Abstract

Under intracellular environment, proteins experience macromolecular crowding effects including specific/nonspecific interactions with other factors, which tunes their physical properties, e.g. 3D structure, dynamics and conformational stability, and thus biological functions [1]. *In situ* observation of protein behaviours is therefore necessary. In-cell NMR [2] is currently the only approach that can provide structural information of proteins inside cells at atomic resolution.

Protein structure determination by in-cell NMR spectroscopy has been achieved in our laboratory for the first time using *E. coli* cells [3][4]. Extending in-cell NMR to study proteins inside eukaryotic cells has been another issue to be investigated, and was successfully tested recently for insect cells like sf9 [5]. In this study I employed human cultured HeLa cells as the host cells for in-cell NMR studies, since I believe that the knowledge obtained by in-"human"-cell NMR can be directly utilised in the advanced medical sciences and the pharmaceutical applications.

I selected human calbindin D_{9k}(P47M+C80) as the model system. For in-"HeLa"-cell NMR experiments, we employed the protocols by Inomata et al.[6]: proteins of interest are overexpressed/labelled in *E. coli*, purified, and efficiently delivered to HeLa cells by linking the proteins with a cell-penetrating peptide (CPP-TAT from HIV-1) covalently. Inside the

HeLa cells the proteins are separated from the CPP-TAT by either endogeneous enzymatic activities or autonomous reductive cleavage. The expression system of calbindin D_{9k}(P47M+C80), which has C-terminal Cys residue for CPP-TAT linking, was prepared. ^{15}N- and ^{13}C/^{15}N-labelled proteins were purified for in-cell NMR experiments. Since a drastic reduction of the experimental time is necessary because of the low abundance of target proteins and short lifetime of the live cell samples, all the in-cell NMR experiments were measured with a nonlinear sampling scheme in the indirect dimension. SOFAST-type pulse techniques were used for extra reduction of the experimental time. All the NMR measurements were performed on a Bruker AVANCE-III 600 spectrometer equipped with a cryogenic probehead.

^{1}H-^{15}N correlation cross peaks due to calbindin D_{9k}(P47M+C80) in HeLa cells were well-resolved, suggesting that the proteins in HeLa cells are properly folded. However, when looking at four conjugative in-cell 2D ^{1}H-^{15}N SOFAST-HMQC experiments (each took 38 mins), some spectral changes - peak disappearance and peak arising over measurements - were noticed. In order to investigate these spectral changes, cross peaks in each in-cell NMR spectrum were analysed by comparing with *in vitro* assignments of calbindin D_{9k}(P47M+C80) in the metal-free (apo), Mg^{2+}- and Ca^{2+}-bound states, which were performed separately. As a result, calbindin D_{9k}(P47M+C80) is in the Mg^{2+}-bound state immediately after the start of NMR experiments, and gradually changed to the Ca^{2+}-bound state, which leads to conformational changes in HeLa cells. In addition, it was

found that this spectral change did not depend on the initial metal-binding state of the CPP-TAT linked proteins, suggesting that the bound metal ions are released during the incorporation process into cells. Incorporated calbindin D_{9k}(P47M+C80) binds Mg^{2+}, since Mg^{2+} is abundant in cells, while Ca^{2+} is controlled at low (~μM) level. Next I introduced extra incubation time (~3 hrs) on culture dish prior to harvesting the cells for the experiments. Essentially identical results suggested that the spectral changes are initiated after the preparation of NMR samples. It is well known that HeLa cells release Ca^{2+} ions under various stresses, and heating caused by NMR radio frequency pulses was thought to be a source. Pre-incubation (~2hr) in NMR tubes inside the magnetic field showed that calbindin D_{9k}(P47M+C80) is in the Ca^{2+}-bound state, suggesting that not only the heating by RF pulses but also other stresses caused by the packing the cells into NMR tubes, e.g. cell precipitation due to gravity and oxygen starvation, stimulate the release of Ca^{2+} in cells. These stresses presumably release Ca^{2+} gradually, which are immediately captured by calbindin D_{9k}(P47M+C80), since its affinity to Ca^{2+} is approximately 10^3 times higher than Mg^{2+}. These results showed that HeLa cells are somehow suffering from various stresses during the in-cell NMR measurements, even though they are alive (the cell viability after 2.5 hrs' NMR experiments is more than 90 %). Further improvements in hardware, e.g. recycling the media so as to provide oxygen and nutrients continuously, would be preferable in order to achieve *in situ* measurements under more physiological condition in future.

References

[1] Ellis, R.J., *Trends. Biochem. Sci.* **26**, 597-604 (2001); [2] Serber, Z. et al., *J. Am. Chem. Soc.* **123**, 2446-2447 (2001); [3] Sakakibara, D. *et al.*, *Nature* **458**, 102-105 (2009); [4] Ikeya, T. et al., *Nat. Protoc.* **5**, 1051-1060 (2010); [5] Hamatsu, J. et al. *J. Am. Chem. Soc.* **135**, 1688–1691 (2013); [6] Inomata, K. *et al.*, *Nature* **458**, 106-109 (2009).

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor Prof. Yutaka Ito for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research, publishing the paper and writing of this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Kouji Hirota and Prof. Masaki Mishima, for their encouragement, insightful comments, and hard questions. My sincere thanks also go to Dr Teppei Ikeya for many discussions during my research period.

A special thanks to Dr Jumpei Hamatsu for his support for my research work and understanding the Japanese life as well.

Many thanks to Dr Tsutomu Mikawa (Cellular and molecular biology unit, RIKEN Advanced Science Institute) for providing human calbindin D_{9k} plasmid and Ms Emiko Seiwa for checking the expression level of calbindin D_{9k}. I would also like to thank Ms Tomomi Hanashima and Ms Kaori Oonishi for their help at the early stage of my doctoral course. Gratitude is also due to all students and staffs of Organic and Structural Biochemistry Laboratory (Department of Chemistry, Tokyo

Metropolitan University) for their help and discussion

I would like to thank Tokyo Metropolitan Government for awarding me the scholarship from Asian Human Resources Fund for my Doctoral study.

Suggestions and helps from Drs Venkat Giri Magupalli and Ravikrishnan Elangovan are really appreciated. Last but not the least; I would like to thank my family: my parents and elder brother for supporting me throughout my life.

Abbreviations

2D	Two Dimensional
3D	Three Dimensional
CaBPs	Calcium Binding Proteins
CPP	Cell Penetrating Peptide
Cryo-EM	Cryo-Electron Microscope
Cryo-ET	cryo-electron tomography
DIPSI	Decoupling In the Presence of Scalar Interaction
DMEM	Dulbeccong In the Presence of Sc
DNA	DeoxyriboNucleic Acid
DTT	DiThio Threito
E. Coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetraacetic Acid
ER	Endoplasmic Reticulum
FID	Free Induction Decay
FM	Forward Maxmum entropy
FRET	Fluorescence Resonance Energy Transfer
HMQC	Heteronuclear Multiple Quantum Coherence
HSQC	Heteronuclear Single Quantum Coherence
IPTG	IsoPropyl β-D-1-ThioGalactopyranoside
K Da	Kilo Dalton

MaxEnt	Maximum Entropy
MDD	MultiDimensional Decompsition
NMR	Nuclear Magnetic Resonance
P. pastoris	*Pichia pastoris*
PBS	Phosphate Buffer Saline
PC9	PolyChromatic 9
QME	Quantitative Maximum Entropy
REBURP	Refocusing Band-Selective Pulse with Uniform Response and Phase
RNA	RiboNucleic acid
SLO	StreptoLysine O
SNOB	Selective excitatioN fOr Biochemical application
SOFAST	band-Selective Optimized Flip Angle Short Transient
TPPI	Time-Proportional Phase Incrementation
Tris	Tris-(hydroxymethyl)aminomethane

Contents

Chapter 1

Macromolecular crowding inside living cells

Intracellular conditions of the living cells are highly crowded by their cytoplasmic macromolecules like proteins, nucleic acids, etc. as well as many organelles like mitochondria, endoplasmic reticulum, golgi complex etc. [1]. Such condition is termed as “crowded” rather than “concentrated”, as no single macromolecular species exist at high concentration, but together. These macromolecules occupy a significant fraction of the total volume of the cytoplasmic space. Biological macromolecules have evolved to function inside such crowded environments. The crowding effects can make molecules in cells behave in different ways than in test-tube assays. Consequently, measurements of the properties of biological macromolecules or processes in biological activities that are made in the laboratory in dilute solutions (*in vitro*) may be different by many orders of magnitude from the true values seen in living cells (*in vivo*). Studies of biochemical processes under realistically

crowded conditions are very important, since the intracellular conditions are a ubiquitous property of all cells and crowding may be essential for the efficient operation of biological phenomena.

The total concentration of proteins and RNAs inside a cell of *Escherichia coli* is in the range of 300–400 g l^{-1} [1]. In eukaryotes, the interior of the cells is further crowded by the protein filaments that make up the cytoskeleton, this meshwork divide the cytosol into a network of narrow pores. Fig. 1.1 shows an artistic view of the interior of a eukaryotic cell.

As was mentioned above, these highly crowded macromolecules occupy a large proportion of the volume of the cells, which reduces the volume of solvent that is available for other macromolecules. This effect, defined as "*excluded volume effect"* [2], has been thought to select the most compact native structures of proteins. Experiments using organic polymers etc. as "crowders" support this theory [3-6]. On the other hand, various non-specific interactions are thought to select rather extended conformations with wider surface area. Therefore, the protein structures and dynamics at intracellular environment exist under a fine balance of these effects (Fig.1.2).

Since it is difficult to mimic the intracellular environment *in vitro*, *in situ* observations become very important. For an example, protein aggregation, binding to substrates and metal upon signalling are the consequences of crowding effects which can’t be studied in *in vitro*.

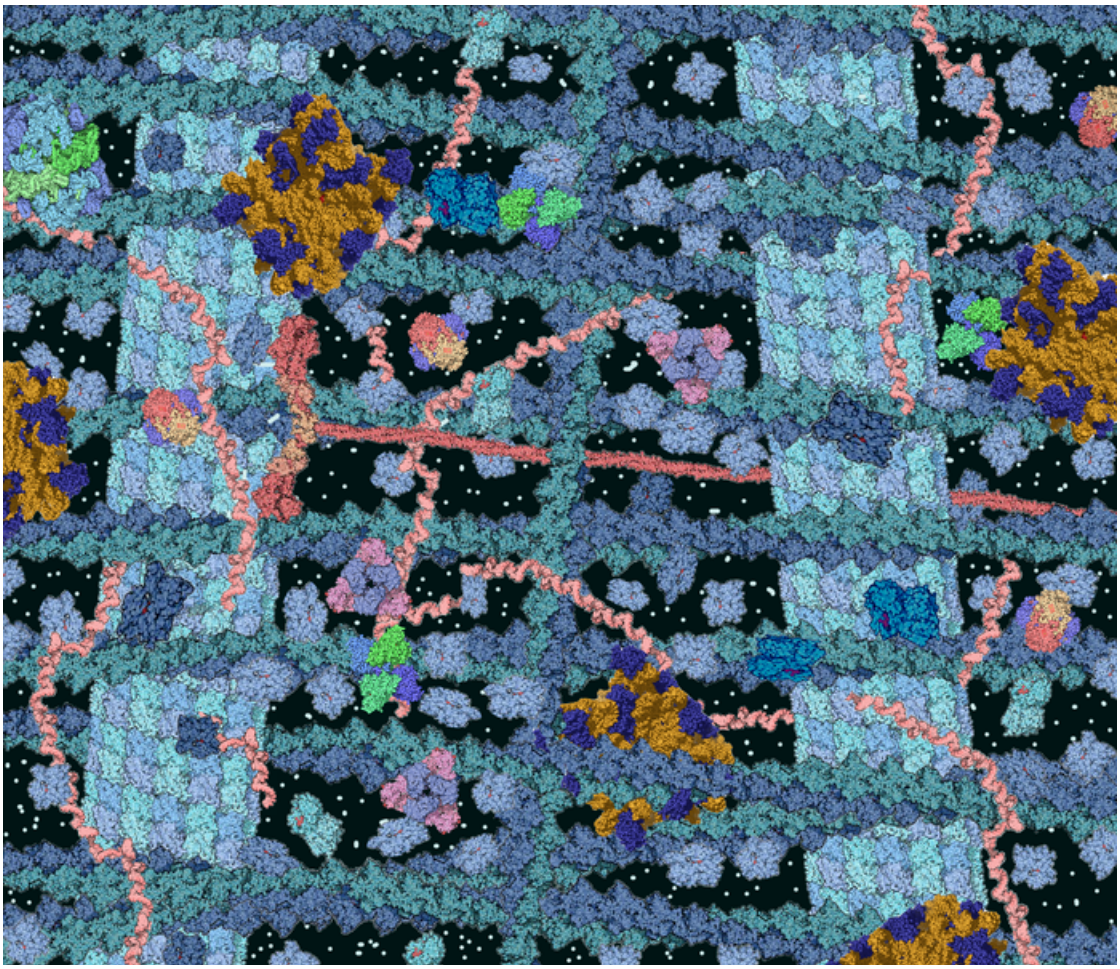


Fig.1.1

An artistic view of macromolecular crowding inside the cell. This figure was downloaded from http://en.wikipedia.org/wiki/Macromolecular_crowding

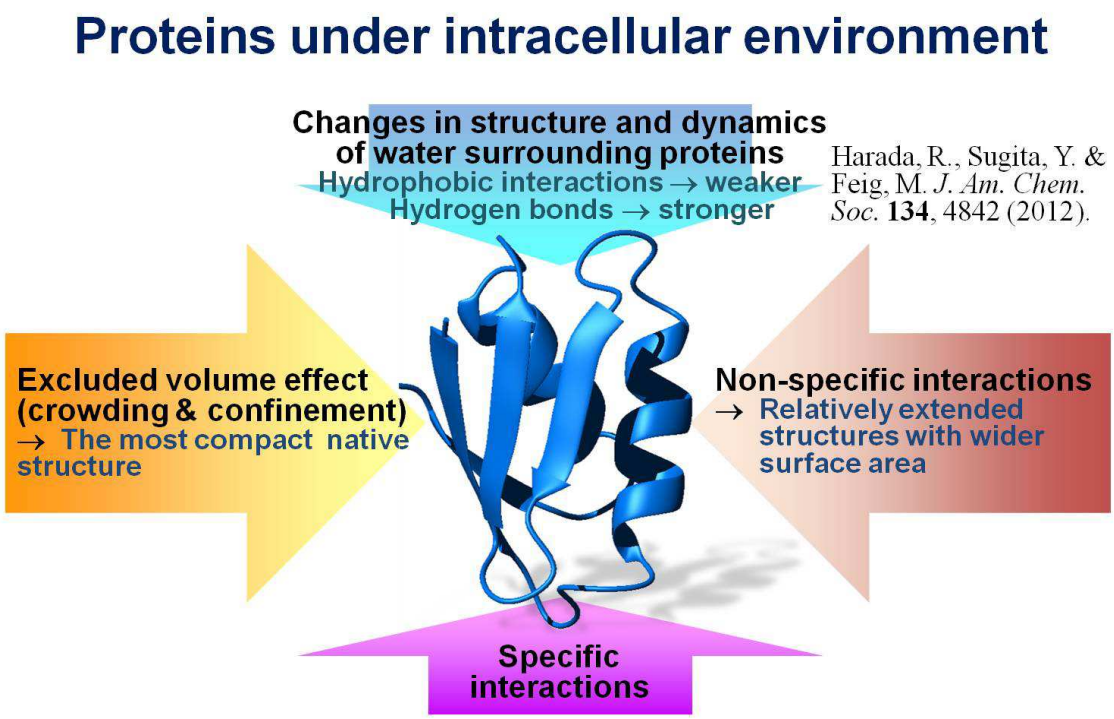


Fig. 1.2

A schematic drawing of the various intracellular effects which may perturb protein structure and dynamics. The changes in structure and dynamics of water surrounding proteins was proposed recently by molecular dynamics simulations [7].

Chapter 2

Methods to study of biomolecules inside living cells

Direct visualization of biomolecules like DNA, RNA, proteins and their structures and biological functions, became a necessary to understand cellular phenomena. Hence analytical techniques were developed to visualize and interpret them. A number of techniques are discussed below.

2.1 Cryo- electron microscope

Cryo-electron microscope (cryo-EM) has been used as a standard technique to study biomolecules in *in vitro* and *in vivo*. Owing to the high magnification that can be achieved, transmission electron microscope (TEM) has helped uncovering the fine architecture of intracellular components and improved our understanding of the physiological process

of cells, in a number of different systems ranging from bacteriophages to mammalian cells [8]. Electron microscope can bridge the gap between ultra-structural compartmentalization and the structural analysis of molecular inhabitants of the cells, at a nanometer-length scale by providing high resolution 3D images (tomograms) of cellular landscapes. The combination of electron microscope with cryo-techniques, termed cryo-electron tomography (cryo-ET), enables the visualization of frozen hydrated biological samples unadulterated by harmful preparation methods. Some of the disadvantages are that it is necessary to keep the cells in vacuum and frizzed, and to construct unadulterated samples by harmful preparation methods, which are exactly not physiological conditions [9].

2.2 Fluorescence resonance energy transfer microscopy

When light microscopy initiated our understanding of cellular structure and the associated function, molecular biological studies over the past few decades have shown that cellular events, such as signal transduction and gene transcription, require the assembly of proteins into specific macromolecular complexes. Traditional biophysical or biochemical methods did not provide direct access to the interactions of these protein partners in their natural environment. Intensity-based imaging techniques applying the method of fluorescence resonance energy transfer (FRET) microscopy was subsequently developed, facilitating the study of

these interactions inside intact living cells [10]. New imaging technologies, coupled with the development of new genetically encoded fluorescent labels and sensors and the increasing capability of computer software for image acquisition and analysis, have enabled more sophisticated studies of protein functions and processes ranging from gene expression to second-messenger cascades and intercellular signalling [11].

Fluorescence resonance energy transfer (FRET) is a distance dependent physical process by which energy is transferred non-radiatively from an excited molecular fluorophore (the donor) to another fluorophore (the acceptor) by means of intermolecular long-range dipole–dipole coupling. FRET can be an accurate measurement of molecular proximity at angstrom distances (10–100 Å) and highly efficient if the donor and acceptor are positioned within the Förster radius (the distance at which half the excitation energy of the donor is transferred to the acceptor, typically 3–6 nm). The efficiency of FRET is dependent on the inverse sixth power of intermolecular separation, making it a sensitive technique for investigating a variety of biological phenomena that produce changes in molecular proximity. An example, capturing Ca^{2+} signaling in specific cellular compartments like the cytoplasm, nucleus, or endoplasmic reticulum can be observed by measuring the change in the ratio of the fluorescence intensities of acceptor and donor molecules in live cells [12]. Cameleons, a class of fluorescent indicators for Ca^{2+} based on GFPs and calmodulin (CAM), are useful tools in measuring the free Ca^{2+} concentrations in living cells.

A disadvantage is, a fluorescence protein as a probe which must be fused to the target protein is needed for FRET based imaging technique.

2.3 In-Cell NMR Spectroscopy

Since its discovery in 1945 the nuclear magnetic resonance (NMR) effect has developed from an interesting physical and chemical phenomenon into the most important analytical technique in chemistry, biology and medicine. Although, originally, NMR seemed to be too insensitive for applications to biological systems, many different techniques based on the NMR phenomenon have emerged during the last 30 years. Today NMR spectroscopy is used for investigations of structure and dynamics of macromolecules as well as binding studies. One aspect that distinguishes NMR from other biophysical techniques is that it can be applied not only to purified *in vitro* samples of targeted macromolecule but also to investigations of living cells as well as entire organisms. In particular, imaging techniques have had and continue to have an enormous impact as a tool in medicine. Imaging techniques use differences in the properties of protons between different tissues to create images. Characteristics that can be used include differences in the T_1 relaxation rates, or in the rate of diffusion. In addition to protons other nuclei have also been used for imaging purposes, for example ^{13}C, ^{10}B, and ^{11}B. While imaging techniques provide information about entire organs, NMR can also

be used to obtain detailed information about molecular processes in living cells and organisms. *In vivo* NMR spectroscopy has for example been used to investigate metabolic fluxes [13-15]. In these experiments small molecules which are labelled with NMR active isotopes, for example ^{13}C, are added to cells. Through the cellular metabolism these NMR active spins get incorporated into other molecules, which can be identified based on their characteristic chemical shifts.

The investigations of molecules in living cellular systems were limited to these classical *in vivo* (and metabolomics) NMR experiments that focus on small molecules. The gap between *in vivo* NMR experiments with small metabolic molecules together with imaging techniques providing information about entire organs and organisms and the behaviour of biological macromolecules inside living cells, bridged by in-cell NMR spectroscopy allowed investigator to understand cellular functions [16]. The observation of biological macromolecules in cellular systems is based on labelling schemes that enable the selective identification of these macromolecules in an environment that is crowded with other macromolecules as well as with many different small molecules. Basically two different approaches to achieve this goal exist. The first is overexpression of isotopically labelled protein inside host cells (*E.coli* [17], yeast [18], Mammalian [19]), and the second is either injecting the isotopically labelled proteins to the host cells in micro level (*Xenopus* Oocyte [20,21]) or incorporation by cell penetrating peptides [22] or pore-forming toxins [23]. These in-cell NMR methods have opened the

way to understand the *in situ* cellular functions and behaviours by looking at atomic level.

Chapter 3

In-cell NMR Spectroscopy

3.1 High resolution multi-dimensional NMR in living cells

On 2001, Serber et.al. established a novel biomolecular NMR technique, so-called “In-cell NMR”, to study protein behaviour at atomic resolution inside living *E. coli* cells [17]. They also proposed potential applications of this approach for investigating various biological events [16] (Fig.3.2.1). It is one of experimental methodologies to study biological macromolecules, such as proteins and DNAs, taking advantage of non-invasive property and atomic resolution of NMR spectroscopy, and uses living cells as NMR sample, in which ^{15}N or $^{13}C/^{15}N$ labelled target macromolecules were incorporated. Since NMR signals of target molecules were directly observed from inside of the cells, behaviour of macromolecules in living cells are feasible to study with this technique. Over the past decade, this method has gradually being developed and

mainly used for analysing the protein behaviours in living prokaryotic and eukaryotic cells [24-26].

3. 2 In-cell NMR using prokaryotic cells.

All of reported in-cell NMR using prokaryotic cells employed *E. coli* cells as “host cell” [25]. *E. coli* cells which were used as in-cell NMR sample overexpressed recombinant target proteins and were labelled with NMR active isotope (Fig.3.2.2). In-cell NMR was first applied to proteins inside *E. coli* cells [17,27]. This method has been used to detect various intracellular behaviours of proteins such as conformational changes [28], dynamics [29], protein-protein [30] and protein-ligand [31] interactions and high resolution 3D structure [32,33] in living bacterial cells. Recently, bacterial in-cell NMR has been often used for investigating macromolecular crowding effect and new insights about protein dynamics [29], structural stability [34] and diffusion behaviours [35] in the crowded environment were reported.

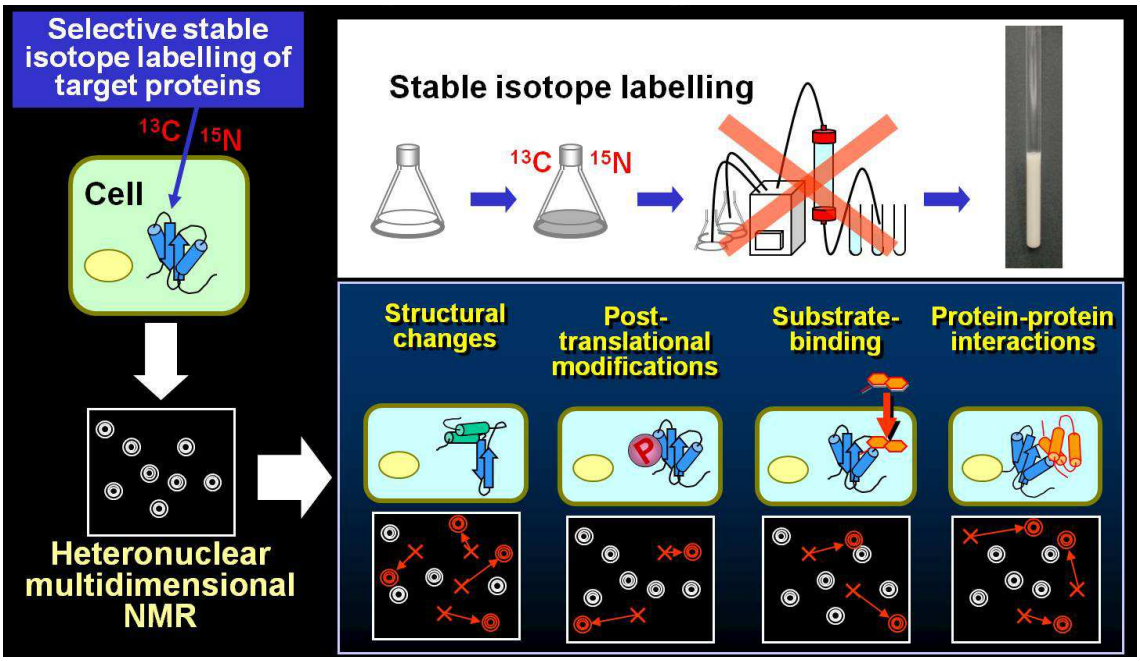


Figure 3.2.1

Potential applications of in-cell NMR to investigate various biological events inside cells.

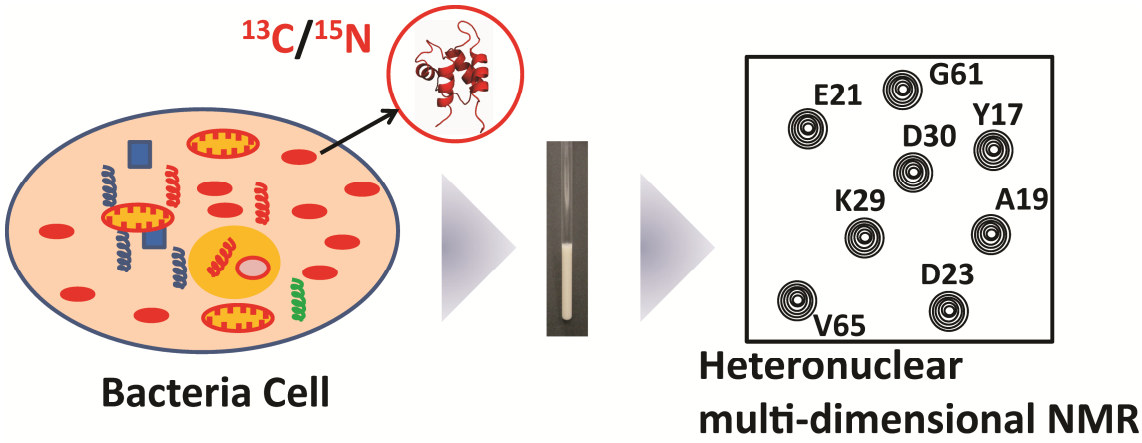


Figure 3.2.2

Schematic illustration of the in-cell NMR experiment in bacteria cells.

3. 3 In-cell NMR using eukaryotic cells.

3.3.1 Eukaryotic cells (non mammalian)

In eukaryotic cells, in-cell NMR studies were first performed by injecting ^{15}N-labelled proteins into *Xenopus laevis* oocytes or eggs [20,21]. Bertrand et al. reported the observation of in-cell NMR spectra of ubiquitin using the protein expression system in *Pichia pastoris* which provide us new insights for properties of Yeast intracellular vesicles [18]. Fig, 3.3.1.1 shows the schematic representation of these approaches.

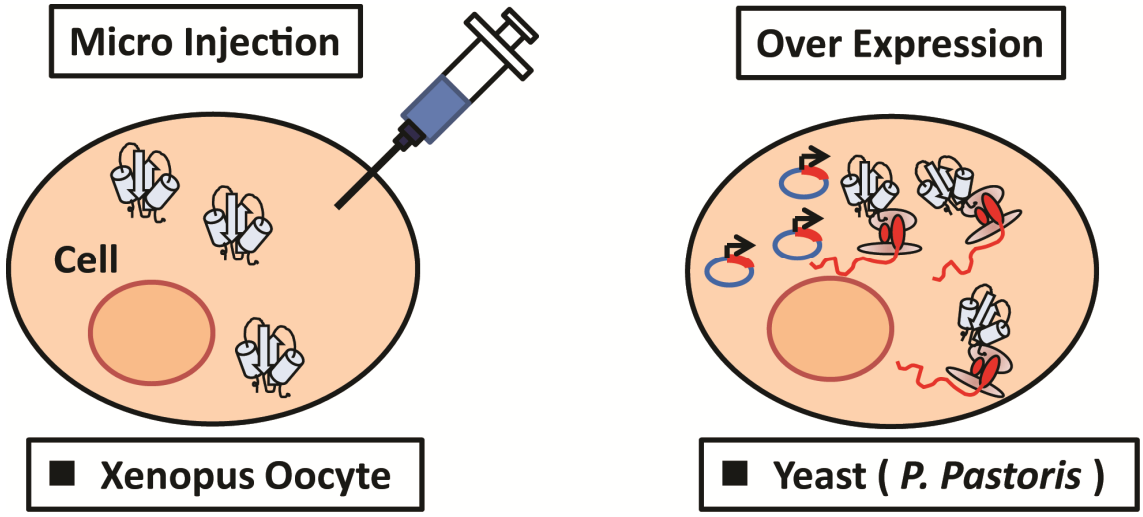


Fig.3.3.1.1

Illustrations of two different kind of sample preparations for in-cell NMR experiments, microinjection of target protein in *Xenopus* Oocyte [20,21] and overexpression of target protein in Yeast (*Pichia pastoris*)[8].

3.3.2 Eukaryotic cells (mammalian)

For cultured mammalian cells, cell-penetrating peptides [22] or the pore-forming toxin streptolysine O (SLO) [23] have been used to deliver proteins. These approaches achieved *in situ* observations of the post-translational modification [23,36] in eukaryotic cells. In addition, applications for pharmaceutical purposes are also anticipated [37] because of the ability of drug screening within living cells. Comparing to the protocols for bacterial in-cell NMR experiments, in which proteins of

interest are overexpressed in host cells, these protocols for eukaryotic cells have the big advantage that the resulting spectra are background free. The existing protocols for introducing stable isotope-labelled proteins into eukaryotic cells however require relatively large quantity of purified and concentrated stable-isotope enriched proteins, thus restricting the applications to proteins which are difficult to purify and/or unstable. The alternative approaches for eukaryotic in-cell NMR utilising the intrinsic protein expression system of host cells have therefore been awaited in order to extending the range of applications. Banci et al. reported about the protein maturation in human cells where proteins were expressed in human cells [19]. Fig.3.3.2.1 shows a schematic representation of these methods described above.

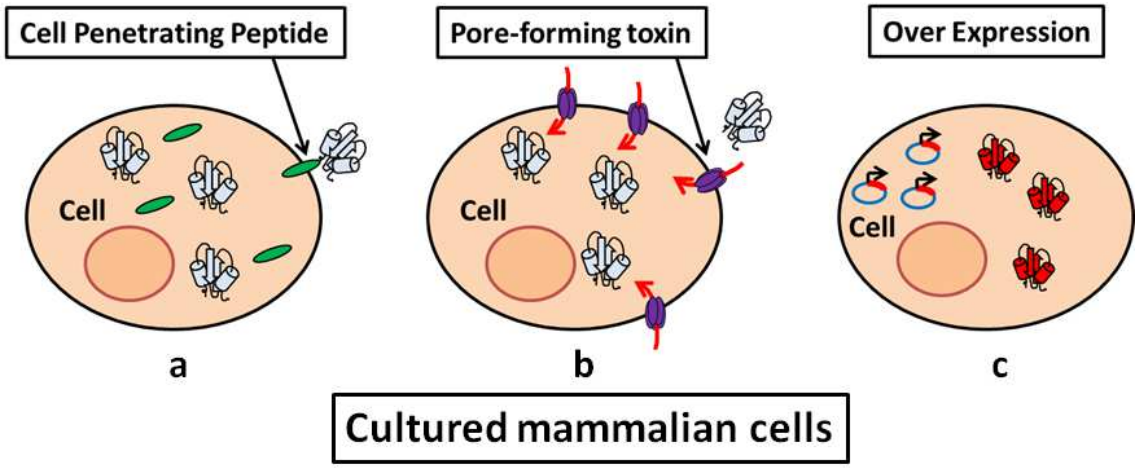


Fig 3.3.2.1

Sample preparations for the culture mammalian cells (a) cell penetrating peptide CPP^{TAT} [22], (b) Pore-forming toxin, Streptolysin O (SLO) [23], (c) Overexpression of target protein [19]

Chapter 4

Aim of the Thesis

In this thesis, I focused on in-cell NMR study of human calbindin D_{9k} inside the cultured human cells (HeLa cells). The aim of the thesis is to extend the in-cell NMR spectroscopic study to proteins in eukaryotic cells, exclusively on HeLa cells. In addition, I have used human calbindin D_{9k} as a target protein to understand its behaviour inside the HeLa cells and Ca^{2+} binding process inside the HeLa cells during the in-cell NMR measurements.

Calcium ion (Ca^{2+}) is the most common signal transduction in living cells, from single cell bacteria to highly developed human cells [38-40]. Many cellular proteins bind Ca^{2+} tightly, in some cases simply to buffer or lower free Ca^{2+} and some trigger second-messenger pathway. Normal intracellular Ca^{2+} level is kept at lower concentration around ~100 nM, since it precipitate phosphate and, thought Ca^{2+}-dependent signal transduction pathways, Ca^{2+} controls a wide variety of cellular process such as cell-cycle progression, differentiation, muscle contraction, enzymatic activities and extracellularly, Ca^{2+} wave and cell morphology [41]. Excess

intracellular Ca^{2+} can result enzymatic breakdown of proteins and cell death by stress. Influx of Ca^{2+} upon stimulation either from intracellular stores or through different types of Ca^{2+} channel leads to an increase in intracellular Ca^{2+} concentration, which eventually triggers apoptosis [42,43]. To maintain low concentration, Ca^{2+} is actively pumped from the cytosol to the extracellular space and into the endoplasmic reticulum (ER), and sometimes in the mitochondria [44]. This allows Calcium binding proteins (CaBPs) to bind Ca^{2+}, undergo a conformational change and associate with different target proteins, hence biological effects of the Ca^{2+} signals.

Calcium binding proteins are usually low molecular weight and acidic protein which are composed of two distinct EF hands flanked by hydrophobic regions at either terminus and separated by a central hinge region. It is believed that this Ca^{2+} binding process is of cooperative way [45]. In contrast to Ca^{2+}, Mg^{2+} is also essential element with structural and catalytic activities, which is quite abundant divalent metal ion within cells (0.5-2.0 $X10^{-3}$ M). In resting cell, the free Ca^{2+} concentration is 1000 fold less than free Mg^{2+}, calcium binding proteins are of Mg^{2+} bound form. If the balance between Mg^{2+} and Ca^{2+} concentration inside cell is disturbed, a semi-stressed state is experienced. As Mg^{2+} is important for stress-response, low level Mg^{2+} concentration means, the stress response can’t subside as it would be. With the increase of Ca^{2+} inside cells due to stress or other biological activities, these CaBPs bind Ca^{2+} and release Mg^{2+}. It is reported that prolonged stress force ER to release calcium and induce cell death and

aging related diseases [46]. This excess calcium should be absorbed by CaBPs in order to prevent from cell death or disorderedness.

Inorder to understand these intracellular phenomena, we initiated in-cell NMR methodology, where human calbindin D_{9k}(P47M+C80) was used as a probe molecule inside human (HeLa) cells (chapter-6) . For the reference and to check the quality of in-HeLa-cell spectra, *in vitro* NMR experiments were performed. (chapter-5). *In vitro* NMR experiments were performed on lysine selective labelled human calbindin D_{9k} of its various form (metal free, Mg2+ and Ca2+ bound form) in order to reduce the ambiguity in interpreting the in-HeLa-cell spectra. Various modifications of the in-Hela-cell sample preparations were carried out to confirm the cellular stress and understanding the intracellular phenomena during the in-Hela-cell NMR experiments (chapter-7). I believe that our result can be used as a reference for understanding the cellular stress and its related intra cellular phenomena by in-cell NMR spectroscopy. Healthiness of the cells during the in-cell NMR measurement is also important information can be obtained easily by our experimental procedure.

Chapter 5

In vitro NMR spectroscopy

5.1 Model System

Human calbindin D_{9k} is a small (79 residue) Ca^{2+} binding protein of low molecular weight (9 K), highly expressed in cytoplasm of mammalian intestinal epithelial cells [41]. It can also be found in the kidney and uterus in some mammalian species [47]. This protein is a member of the S100 superfamily of calcium binding proteins. It has two EF-hands sequences which bind Ca^{2+} with high affinity [48]. Calbindin D_{9k} acts like a buffer protein to control the Ca^{2+} concentration [49] and has less interaction partner in cytoplasm, makes it as a suitable candidate for our study.

5.2 Sample preparation for *in vitro* NMR spectroscopy

5.2.1 Expression

The DNA fragment encoding the human calbindin D_{9k} gene containing the proline 47 to methionine (P47M) mutation and the C-terminal additional cysteine residue for the CPP^{TAT} conjugation [henceforth referred as calbindin D_{9k}(P47M+C80)] was synthesised. The P47M mutation was introduced so as to restrict *cis-trans* isomerisation [50]. The DNA fragment was inserted into a pET3a vector for overexpression in the *E.coli* BL21 Star (DE3) pLysS strain (Invitrogen). Uniformly $^{13}C/^{15}N$-labelled samples were prepared by growing the transformed bacteria in M9 minimal medium containing $^{15}NH_4Cl$ and [$^{13}C_6$]-D-glucose as the sole sources of nitrogen and carbon, respectively. For the preparation of uniformly ^{15}N-labelled samples, [$^{13}C_6$]-D-glucose was replaced by the same concentration of unlabelled D-glucose. The lysine-selectively ^{15}N-labelled samples were prepared by using M9 minimal medium containing 100 mg/L ^{15}N-lysine. Cells were grown at 37 °C, and protein expression was induced by the addition of isopropyl thio-β-D-thiogalactoside. After ~14 hours of further growth, the cells were harvested.

5.2.2 Purification of human calbindin D_{9k}(P47M+C80)

All the following steps in the purification protocol were carried out

at 4 °C unless mentioned. Re-suspending in the lysis buffer [20 mM imidazole (pH 7.0), 20 mM NaCl and 1 mM EDTA], the harvested cells were lysed by sonication. The cleared lysate was prepared by ultra-centrifugation, and then mixed with equal amount of the pre-heated (95 °C) lysis buffer, and kept at 95 °C for 3 minutes. After cooling down on ice, precipitated proteins were removed by another ultra-centrifugation. The supernatant was then loaded onto a DEAE-Sepharose (GE Healthcare) column pre-equilibrated with the lysis buffer. The column was washed with 5 column volume of the same buffer, followed by elution with a linear gradient of 20-300 mM NaCl. The fractions containing calbindin D_{9k}(P47M+C80) were loaded onto a Superdex-75 16/60 column (GE Healthcare) pre-equilibrated with the preparation buffer [20 mM Tris-HCl (pH 7.4), 20 mM NaCl , and 1 mM EDTA]. Further purification was carried out by Resource-Q (GE Healthcare) column chromatography with a linear gradient of 0-200 mM NaCl.

The purified calbindin D_{9k}(P47M+C80), which was in the metal-free state, was concentrated (~1.5 mM) and dissolved in the NMR buffer [20 mM Tris-HCl (pH7.4), 150 mM NaCl, and 5 mM DTT]. For the preparation of samples in the Mg^{2+}- and Ca^{2+}-bound states, $MgCl_2$ and $CaCl_2$ solutions were added at the final concentration of 5 mM, respectively

5.3 Heteronuclear 3D NMR experiments

High resolution multi-dimensional NMR [51,52] is widely used for the field of protein sciences because of its capability to observe accurate structural and dynamic information of proteins in atomic resolution. When focusing on protein structural studies, X-ray crystallography is more popular and powerful tool. One of notable advantages of protein NMR spectroscopy is, however, that it has ability to analyse high resolution protein structures and dynamics in a solution state. Since most of the proteins work in a solution state in living cells, NMR spectroscopy can only provide detailed information of protein behaviours under the condition close to the physiological one.

In the last 30 years, solution state NMR techniques for studies of protein structures and dynamics have been drastically developed. It is well known that one of the major bottle-neck of the protein structure determination by NMR is the molecular size and therefore many methods in NMR spectroscopy were developed to analyse higher molecular weight proteins.

Structural information of target molecules obtained by solution NMR contain chemical shifts, scalar coupling constants, (residual) dipolar couplings, line shapes and intensity in NMR spectra. To link appropriate atoms of proteins with NMR derived information, chemical shift assignments are the most important task for protein analysis. Nowadays unambiguous assignments of backbone and side-chain NMR-observable

atoms are performed by the combination of 3D/4D triple-resonance NMR [53] spectra which provide intra- and inter-residual connectivities of observed resonances.

On the other hand, since tens of thousands of FIDs have to be measured, a considerable drawback of high dimensional (3D/4D) NMR experiments is that it is necessity to spend long measurement time (~ days for a set of 3D experiments/ ~ weeks for a set of 4D experiments). It implies that unstable proteins are not suitable for these experiments. To solve this problem, rapid multi-dimensional NMR measurement techniques such as Non-linear sampling [54,55], SOFAST [56,57], and Single-scan NMR [58], and so on, were developed. Drastic reduction of experimental time with high resolution spectra can be observed using these techniques.

5.4 Non-linear sampling and Data Processing

As was described in the previous section, one of the major draw-back of multidimensional NMR is a lot of experimental time needed for observing 3D/4D spectra. Total number of FIDs increased depending on the number of data points in indirect dimensions. Number of scans for each FID and the recycling delay time also increase the total measurement time.

The non-linear sampling scheme [54] utilises the same pulse schemes of conventional multi-dimensional NMR experiments, but the FIDs are acquired in a non-linear fashion. Since the FIDs are acquired in

non-linear fashion, discrete Fourier transformation, which is widely used for NMR data processing, is not applicable for the indirect dimensions any more. To obtain analysable spectra from the non-linearly sample data, a variety of mathematical processing techniques such as Maximum Entropy (MaxEnt) [59,60], Multi-Dimensional Decomposition (MDD) [61,62], Forward MaxEnt (FM) [63], Compressed Sensing (CS) [64,65], and quantitative Maximum Entropy (QME) [66] etc. have been reported. These techniques provide multi-dimensional spectra with much higher resolution from the data acquired with non-linearly sampling schemes when comparing to the data with the same number of FIDs, acquired with conventional sampling scheme.

5.5 Heteronuclear NMR experiments for backbone resonance assignments

5.5.1 2D ^{1}H-^{15}N HSQC.

2D ^{1}H-^{15}N HSQC spectrum was recorded on ^{15}N-labelled samples with 8 transients. A total of 1024 (t_2, $^{1}H^{N}$) × 256 (t_1, ^{15}N) complex points were acquired for *in vitro* sample.

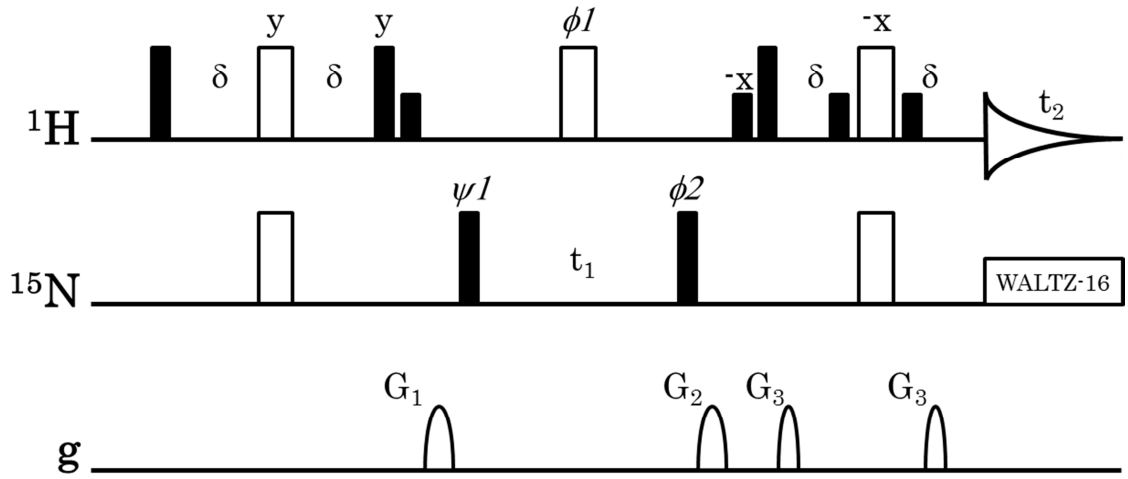


Fig. 5.5.1.1

Pulse sequence for the 2D ^{1}H-^{15}N HSQC experiment. Narrow filled-in and wide open bars represent 90° and 180° pulses, respectively. Short filled-in bars indicate ~1.4 ms 90° water flip-back pulses. Unless indicated, pulses are applied on the x axis. The G_1, G_2 and G_3 gradient pulses are smoothed-square shaped and have durations of 2.3, 0.8 and 0.4 ms, respectively. Their corresponding peak powers are 10, -10 and -50 G/cm along the z-axis. Delay: δ = 2.3 ms. Phase cycling: $\psi 1$ = x, -x; $\phi 1$ = 4(x), 4(-x); $\phi 2$ = 2(x), 2(-x); Rec. = x, -x, -x, x. The phase sensitive detection in ^{15}N dimension is achieved by incrementing phase $\psi 1$ according to States-TPPI scheme [67].

5.5.2 3D CBCA(CO)NH/CBCANH

To observe inter- and intra-residual correlations and chemical shifts of $^{13}C^{\alpha}$ and $^{13}C^{\beta}$, 3D CBCA(CO)NH [68] and CBCANH [69] were measured. The 3D CBCA(CO)NH and CBCANH spectra were recorded on $^{13}C/^{15}N$-labelled samples with 8 transients and 512 complex points for the acquisition dimension (t_3, $^{1}H^{N}$). For indirect dimension, a total 264 of FIDs from 48 (t_1, $^{13}C^{\alpha/\beta}$) × 22 (t_2, ^{15}N) sampling space were randomly acquired. In both experiments, the WATERGATE block [70] was implemented before acquisition in the pulse sequences for solvent suppression.

5.5.3 3D HNCO/HN(CA)CO

To observe inter- and intra-residual correlation and chemical shifts of $^{13}C'$, 3D HNCO [71] and HN(CA)CO [72] were also measured. The 3D HNCO and HN(CA)CO spectra were recorded on $^{13}C/^{15}N$-labelled samples with 8 transients and 512 complex points for the acquisition dimension (t_3, $^{1}H^{N}$). For indirect dimension, a total 352 of FIDs from 64 (t_1, $^{13}C'$) × 22 (t_2, ^{15}N) sampling space were randomly acquired. In both experiments, the WATERGATE block was implemented before acquisition in the pulse sequences for solvent suppression.

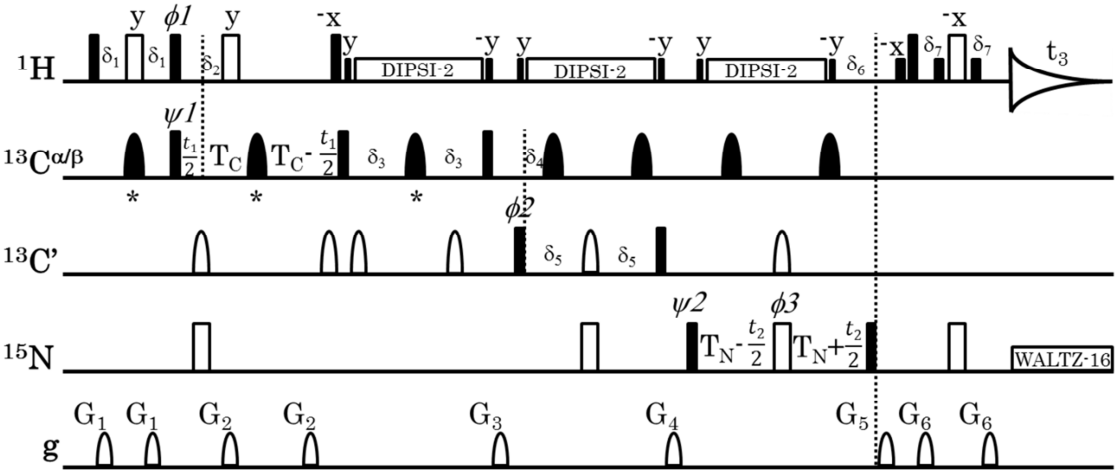


Fig.5.5.2.1

Pulse sequence for the 3D CBCA(CO)NH experiment. Narrow filled-in and wide open bars represent 90° and 180° pulses, respectively. Short filled-in bars indicate ~1.4 ms 90° water flip-back pulses. The 90° pulses flanking DIPSI-2 [73] decoupling intervals which are applied using a ~4.2 kHz field. ^{1}H DIPSI-2 decoupling is achieved using a ~4.2 kHz field. Rectangle 90° pulses for $^{13}C^{\alpha/\beta}$ and ^{13}C' are applied using 5.1 and 4.7 kHz field, respectively. Filled-in shaped pulses marked with " * " have an Q374 profile with duration 256 ms and $^{13}C^{\alpha/\beta}$ carrier is positioned at 40 ppm, while without "*" indicate have RSNOB [75] profile with 367 ms duration and applied at 52 ppm for $^{13}C^{\alpha}$. Open shaped pulses have SEDUCE-1[76] profile with 192 ms and applied at 173 ppm. Unless indicated, pulses are applied on the x axis. The G_1, G_2, G_3, G_4, G_5 and G_6 gradient pulses are smoothed-square shaped and have durations of 0.4, 0.4, 0.8, 1.8, 2.3 and 0.4 ms, respectively. Their corresponding peak powers are 20, 12, 10, 10, 10 and 80 G/cm along the z-axis. Delays: δ_1 = 1.5 ms, δ_2 = 1.1 ms, δ_3 = 3.8 ms, δ_4 = 4.4 ms, δ_5 = 12.4 ms, δ_6 = 5.4 ms, δ_7 = 2.3 ms, T_C = 3.1 ms, T_N = 12.0 ms. Phase cycling: $\psi 1$ = 8(x), 8(-x); $\psi 2$ = 4(x), 4(-x); $\phi 1$ = y, -y; $\phi 2$ = 2(x), 2(-x); $\phi 3$ = 4(x), 4(-x); Rec. = x, 2(-x), x, -x, 2(x), -x, x, 2(-x), x. Phase sensitive detections in ^{13}C and ^{15}N dimension are achieved by incrementing phases $\psi 1$ and $\psi 2$ according to States-TPPI scheme.

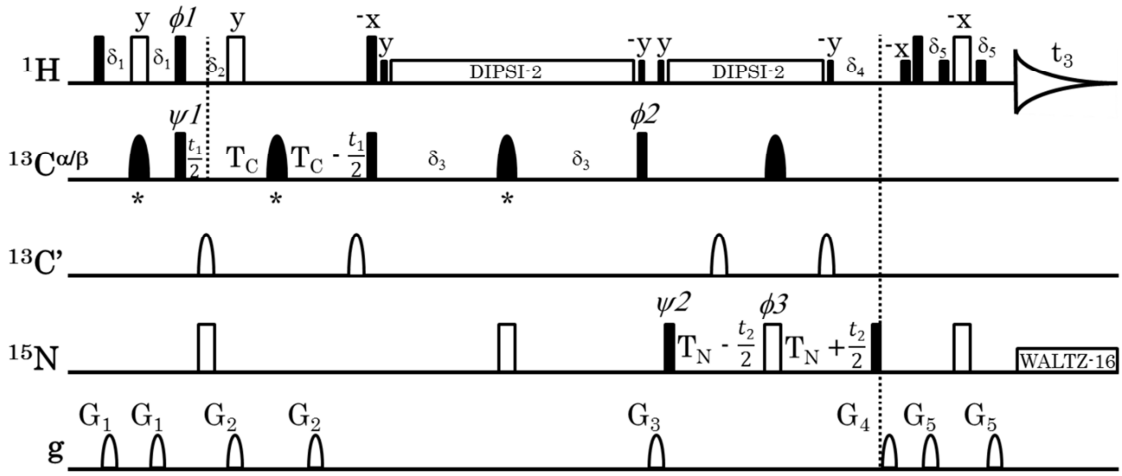


Fig.5.5.2.2

Pulse sequence for the 3D CBCANH experiment. Narrow filled-in and wide open bars represent 90° and 180° pulses, respectively. Short filled-in bars indicate ~1.4 ms 90° water flip-back pulses. The 90° pulses flanking DIPSI-2 decoupling intervals which are applied using a ~4.2 kHz field. ^{1}H DIPSI-2 decoupling is achieved using a ~4.2 kHz field. Rectangle 90° pulse for $^{13}C^{\alpha/\beta}$ is applied using a 5.1 kHz field. Filled-in shaped pulses marked with " * " have a Q3 profile with duration 256 ms and $^{13}C^{\alpha/\beta}$ carrier is positioned at 40 ppm, while without "*" indicate have a RSNOB profile with 367 ms duration and applied at 52 ppm for $^{13}C^{\alpha}$. Open shaped pulses have a SEDUCE-1 profile with 192 ms and applied at 173 ppm. Unless indicated, pulses are applied on the x axis. The G_1, G_2, G_3, G_4 and G_5 gradient pulses are smoothed-square shaped and have durations of 0.4, 0.4, 1.8, 2.3 and 0.4 ms, respectively. Their corresponding peak powers are 20, 12, 10, 10 and 80 G/cm along the z-axis. Delays: δ_1 = 1.5 ms, δ_2 = 1.1 ms, δ_3 = 11.0 ms, δ_4 = 5.3 ms, δ_5 = 2.3 ms, T_C = 3.1 ms, T_N = 12.0 ms. Phase cycling: $\psi 1$ = 8(x), 8(-x); $\psi 2$ = 4(x), 4(-x); $\phi 1$ = y, -y; $\phi 2$ = 2(x), 2(-x); $\phi 3$ = 4(x), 4(-x); Rec. = x, 2(-x), x, -x, 2(x), -x, x, 2(-x), x. Phase sensitive detections in ^{13}C and ^{15}N dimension are achieved by incrementing phases $\psi 1$ and $\psi 2$ according to States-TPPI scheme.

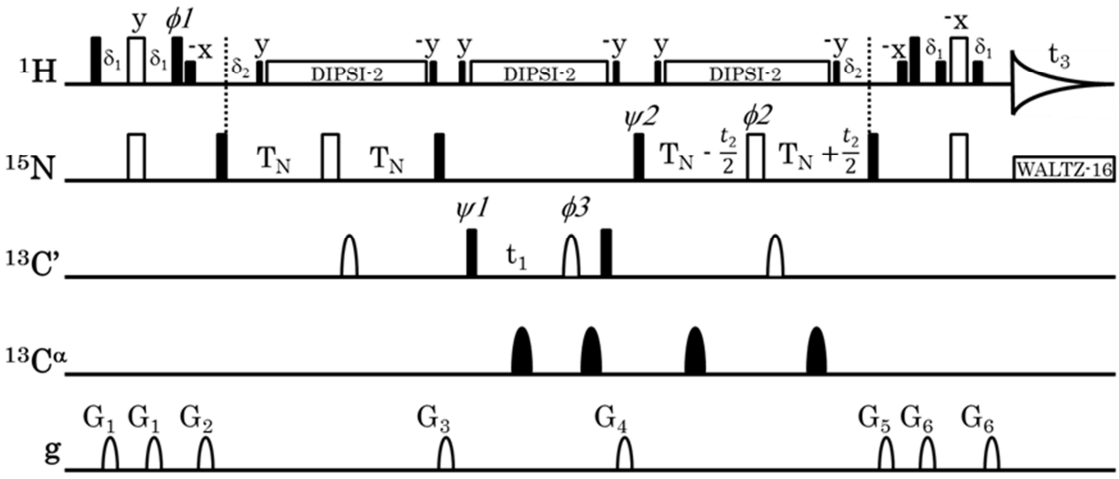


Fig.5.5.3.1

Pulse sequence for the 3D HNCO experiment. Narrow filled-in and wide open bars represent 90° and 180° pulses, respectively. Short filled-in bars indicate ~1.4 ms 90° water flip-back pulses. The 90° pulses flanking DIPSI-2 decoupling intervals which are applied using a ~4.2 kHz field. Rectangle 90° pulse for ^{13}C' is applied using a 4.7 kHz field. ^{1}H DIPSI-2 decoupling is achieved using a ~4.2 kHz field. Filled-in shaped pulses have RSNOB profile with 367 ms duration and applied at 52 ppm for $^{13}C^{\alpha}$. Open shaped pulses have SEDUCE-1 profile with 192 ms and applied at 173 ppm. Unless indicated, pulses are applied on the x axis. The G_1, G_2, G_3, G_4, G_5 and G_6 gradient pulses are smoothed-square shaped and have durations of 0.4, 2.3, 1.8, 1.3, 0.8 and 0.4 ms, respectively. Their corresponding peak powers are 5, 10, 10, 10, 10 and 50 G/cm along the z-axis. Delays: δ_1 = 2.3 ms, δ_2 = 5.5 ms, T_N = 12.0 ms. Phase cycling: $\psi 1$ = 2(x), 2(-x); $\psi 2$ = 4(x), 4(-x); $\phi 1$ = y, -y; $\phi 2$ = 2(x), 2(-x); $\phi 3$ = 4(x), 4(-x); Rec. = x, -x, -x, x, -x, x, x, -x. Phase sensitive detections in ^{13}C and ^{15}N dimension are achieved by incrementing phases $\psi 1$ and $\psi 2$ according to States-TPPI scheme.

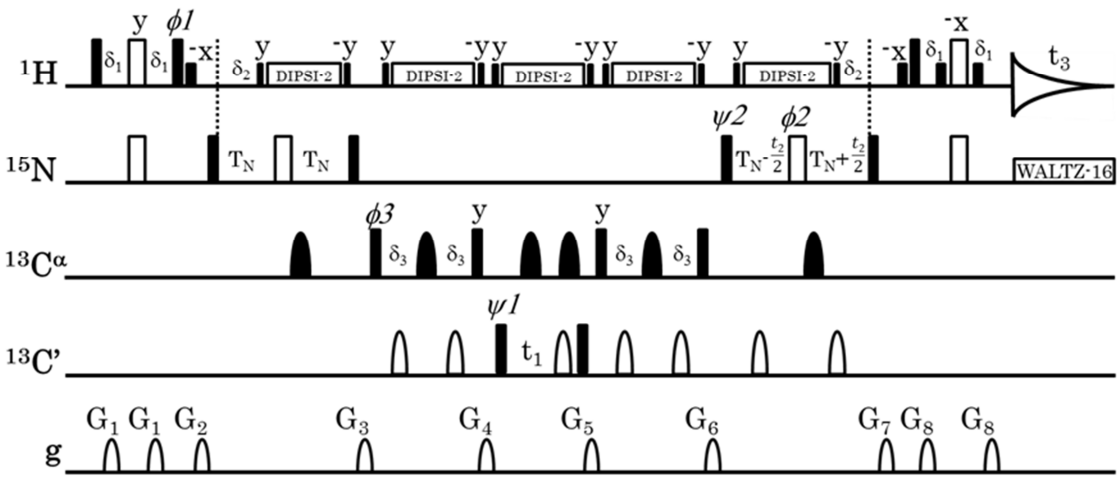


Fig. 5.5.3.2

Pulse sequence for the 3D HN(CA)CO experiment. Narrow filled-in and wide open bars represent 90° and 180° pulses, respectively. Short filled-in bars indicate ~1.4 ms 90° water flip-back pulses. The 90° pulses flanking DIPSI-2 decoupling intervals which are applied using a ~4.2 kHz field. Rectangle 90° pulses for $^{13}C^{\alpha}$ and $^{13}C'$ are applied using a 4.7 kHz field. ^{1}H DIPSI-2 decoupling is achieved using a ~4.2 kHz field. Filled-in shaped pulses have RSNOB profile with 367 ms duration and applied at 52 ppm for $^{13}C^{\alpha}$. Open shaped pulses have SEDUCE-1 profile with 192 ms and applied at 173 ppm. Unless indicated, pulses are applied on the x axis. The G_1, G_2, G_3, G_4, G_5, G_6, G_7 and G_8 gradient pulses are smoothed-square shaped and have durations of 0.4, 2.3, 1.8, 1.5, 1.2, 1.0, 0.8 and 0.4 ms, respectively. Their corresponding peak powers are 5, 10, 10, 10, 10, 10, 10 and 50 G/cm along the z-axis. Delays: δ_1 = 2.3 ms, δ_2 = 5.5 ms, δ_3 = 3.8 ms, T_N = 12.0 ms. Phase cycling: $\psi 1$ = 4(x), 4(-x); $\psi 2$ = 8(x), 8(-x); $\phi 1$ = y, -y; $\phi 2$ = 4(x), 4(-x); $\phi 3$ = 2(x), 2(-x); Rec. = x, 2(-x), x, -x, 2(x), -x, x, 2(-x), x. Phase sensitive detections in ^{13}C and ^{15}N dimension are achieved by incrementing phases $\psi 1$ and $\psi 2$ according to States-TPPI scheme.

5.6 Data processing and analysis

For 2D NMR data, a 1D maximum entropy reconstruction was applied for the indirectly acquired ^{15}N-dimension after processing the directly acquired dimension by Fourier transformation using Azara v2.8 (W. Boucher, www.bio.cam.ac.uk/azara). For 3D NMR data, a 2D quantitative maximum entropy (QME) [66,77] reconstruction was applied for the indirectly acquired dimensions (t1 and t2) after processing the directly acquired dimension (t3) by Fourier transformation.

QME provides improved results in comparison to the maximum entropy (MaxEnt) [59] reconstruction in the Azara v2.8 software. In particular, while one has to arbitrarily determine the Lagrange multipler, λ, in the original MaxEnt recon-struction, in the QME calculation one determines it by seeking the extremum of an approximated conditional probability distribution of the experimental data given λ. Hence, this achieves the optimal λ for an entire spectrum ameliorating some of the problems faced when reconstructing spectra such as those in which there is wide dynamic range, as is often the case in in-cell spectra.

All NMR spectra were analysed using the CcpNmr Analysis 2.2.2 software [78].

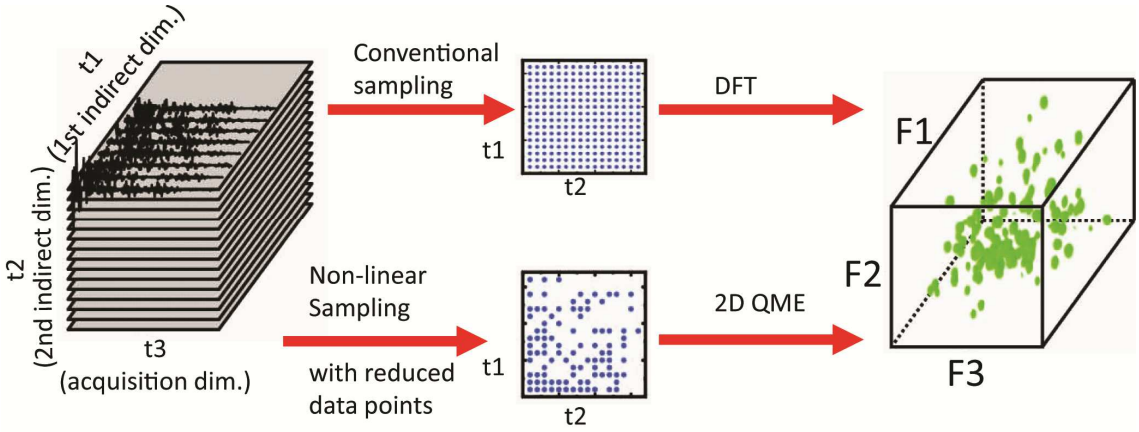


Fig. 5.5.1

A schematic diagram showing the 2D QME processing protocol, by which an equivalent quality of the spectra can be obtained when non-linear sampling scheme is used in contrast to conventional sampling scheme for the reduction of NMR measurement time.

5.7 Backbone resonance assignments

The standard procedures for backbone resonance assignments of small to medium size proteins use 3D CBCANH and 3D CBCA(CO)NH spectra. 3D CBCANH correlates each NH group with C^{α} and C^{β} chemical shift of its own residue (strongly, so peaks are appeared to be strong) and of the residue preceding (weakly, so weak peaks are also appeared). 3D CBCA(CO)NH only correlates the NH group to the preceding C^{α} and C^{β} chemical shifts. Fig. 5.7.1 shows the strategy for backbone resonance assignment of proteins by analysing CBCA(CO)NH and CBCANH spectra.

Supplementing with 3D HN(CA)CO and 3D HNCO spectra, where HN(CA)CO correlates each NH(i) group with C'(i) and C'(i-1), while HNCO correlates each NH(i) group with C'(i-1) (C' referred to carbonyl carbon). Combining all these four 3D triple-resonance spectra and following the assignment strategies, we have performed the backbone resonance assignments for the metal free (Fig 5.7.3), Mg^{2+} bound (Fig 5.7.4) and Ca^{2+} bound (Fig.5.7.5) human calbindin D_{9k}. (results are summarised in Table 5.7.1, 5.7.2, 5.7.3 respectively). Selected ^{1}H-^{13}C strips from the CBCA(CO)NH, CBCANH, HNCO and HN(CO)CA spectra measured on the calbindin D_{9k}(P47M+C80) in the Ca^{2+}-bound state are shown in Fig.5.7.2.

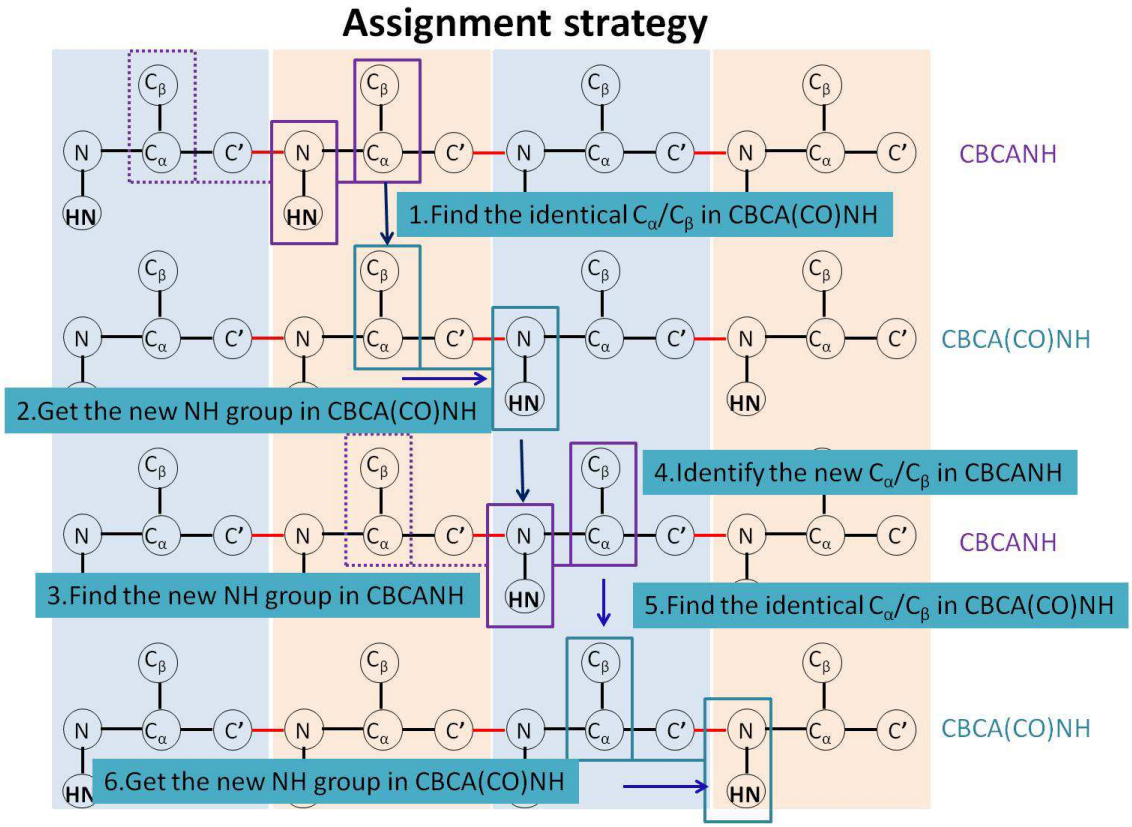


Fig 5.7.1

The strategy for backbone resonance assignment of proteins by analysing 3D CBCA(CO)NH and CBCANH spectra.

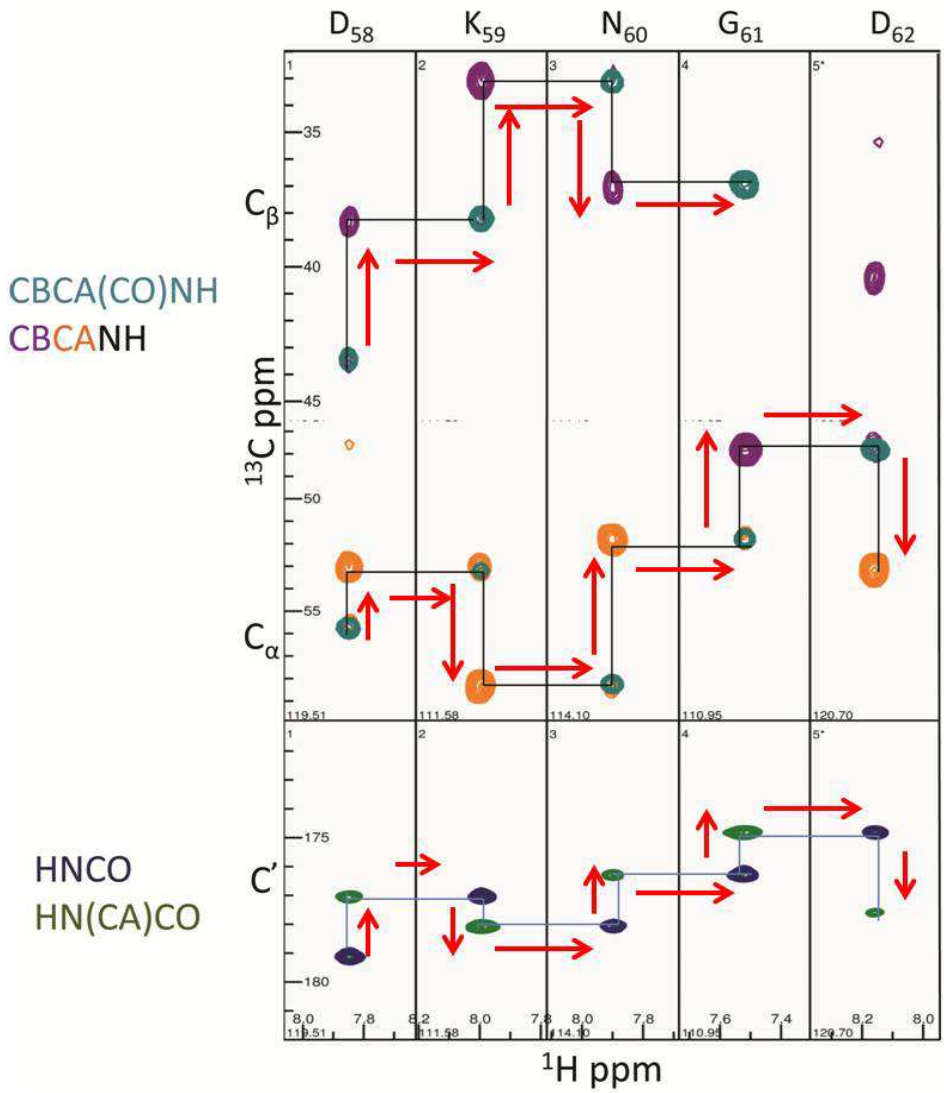


Fig 5.7.2

Selected $^{1}H^{N}$-^{13}C strips extracted from the 3D CBCA(CO)NH and CBCANH spectra (overlaid, upper panel) and the 3D HNCO and HN(CA)CO spectra (overlaid, lower panel) of calbindin D_{9k}(P47M+C80) in the Ca^{2+}-bound state. Each strip corresponds to the ^{15}N frequency of the residue indicated. Sequential connectivities are represented by lines.

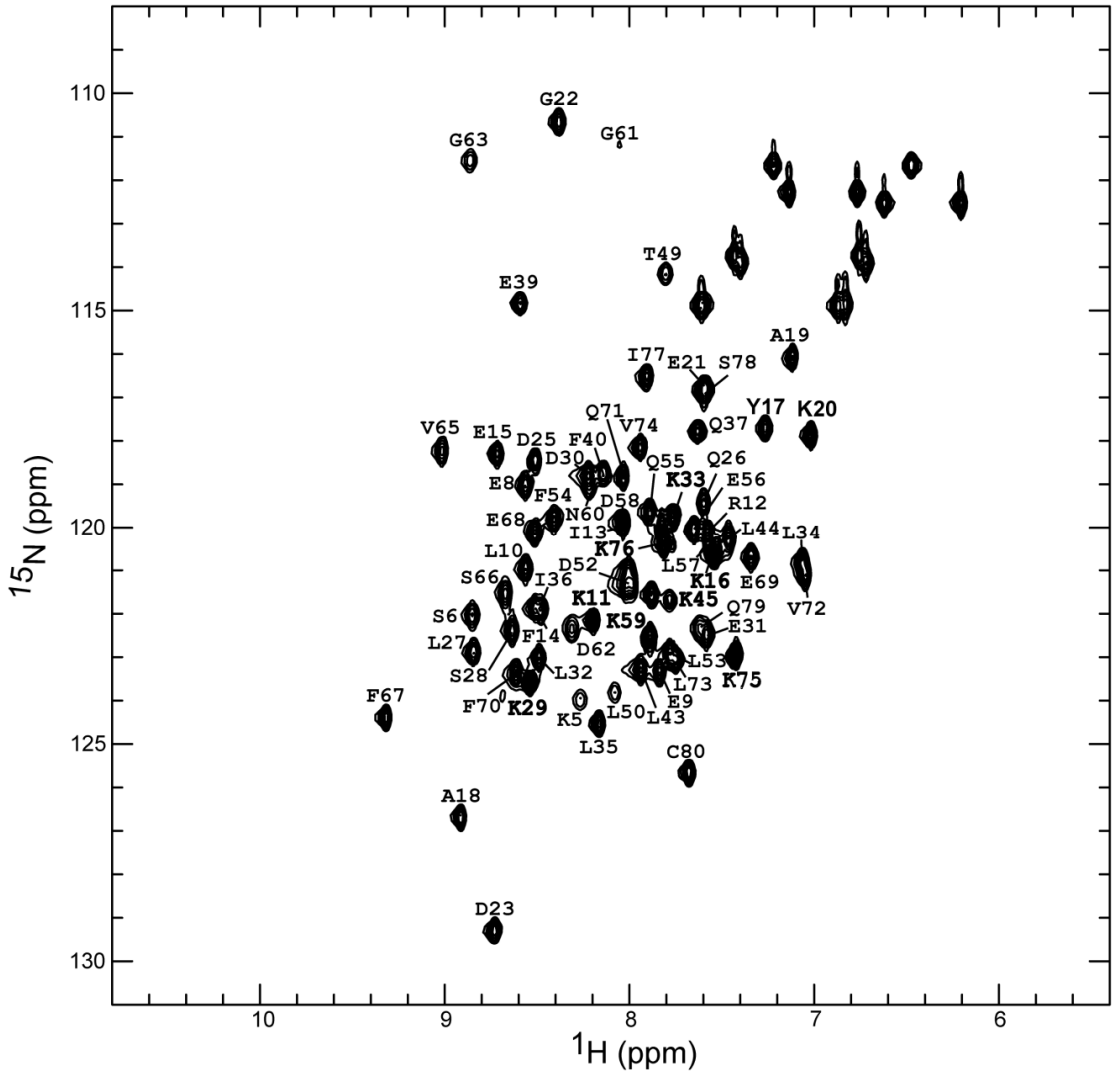


Fig 5.7.3

^{15}N-HSQC showing the backbone resonance assignment of the metal free (Apo) human calbindin D_{9k} (P47M+C80).

Table 5.7.1

Backbone resonance assignment of the metal free human calbindin D_{9k}(P47M+C80) *in vitro* at 37 °C.

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
1	Met	-	-	-	-	-
2	Ser	-	-	-	-	-
3	Thr	-	-	-	-	-
4	Lys	-	-	-	-	-
5	Lys	8.34	124.11	176.12	55.31	34.31
6	Ser	8.84	121.92	173.1	56.79	63.28
7	Pro	-	-	179.17	66.21	31.9
8	Glu	8.55	118.98	179.12	60.17	29.14
9	Glu	7.84	123.18	179.5	59.15	30.18
10	Leu	8.54	120.81	178.83	58.4	42.06
11	Lys	8.2	122.09	177.69	59.58	31.99
12	Arg	7.58	119.97	179.88	59.43	30.42
13	Ile	8.01	121.23	176.77	65.51	38.9
14	Phe	8.5	121.81	176.96	62.87	40.14
15	Glu	8.71	118.25	178.93	59.61	29.9
16	Lys	7.55	120.52	177.92	59.08	32.29
17	Tyr	7.27	117.66	178.15	60.81	39.77
18	Ala	8.9	126.58	178.02	54.77	19.02
19	Ala	7.13	116.13	178.79	53	19.03
20	Lys	7.02	117.82	176.47	59.44	33.05
21	Glu	7.59	116.68	174.98	55.01	32.76
22	Gly	8.37	110.59	174.36	45.95	-
23	Asp	8.69	129.12	176.71	52.16	41.86
24	Pro	-	-	175.16	64.03	31.48
25	Asp	8.5	118.34	175.51	53.48	41.99

Table 5.7.1 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
26	Gln	7.6	119.32	173.35	54.46	33.57
27	Leu	8.83	122.85	175.97	52.97	46.34
28	Ser	8.63	122.29	175.13	57.61	65.12
29	Lys	8.54	123.55	178.15	60.7	31.95
30	Asp	8.22	118.79	179.21	57.63	40.56
31	Glu	7.62	122.48	178.49	58.69	30.04
32	Leu	8.48	122.95	177.97	57.9	41.07
33	Lys	7.76	119.66	178.4	60.95	32.3
34	Leu	7.06	120.7	178.86	58	41.34
35	Leu	8.16	124.44	179.46	59.18	42.29
36	Ile	8.47	121.79	177.4	65.33	37.17
37	Gln	7.63	117.71	176.91	59.14	29.38
38	Ala	7.89	119.65	180.12	54.26	20.2
39	Glu	8.58	114.75	176.75	56.02	30.2
40	Phe	8.13	118.75	173.45	55.01	40.03
41	Pro	-	-	-	65.41	31.21
42	Ser	8.26	115.48	175.78	60.48	62.55
43	Leu	7.92	123.21	177.45	55.83	41.75
44	Leu	7.46	120.11	177.66	55.8	41.57
45	Lys	7.77	121.54	177.19	57.29	32.47
46	Gly	8.09	110.39	-	45.37	-
47	Met	-	-	-	-	-
48	Asn	-	-	175.75	54.7	38.73
49	Thr	7.81	114.06	175.34	62.55	70.06
50	Leu	8.1	123.73	177.73	57.57	41.67
51	Asp	8.02	119.81	178.63	57.88	40.42
52	Asp	7.88	121.4	178.66	57.25	40.5

Table 5.7.1 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
53	Leu	7.76	122.81	178.92	57.58	42.06
54	Phe	8.39	119.67	177.65	60.44	38.29
55	Gln	7.83	119.95	178.16	58.91	28.61
56	Glu	7.66	120	177.93	58.4	29.75
57	Leu	7.55	120.39	177.47	55.85	42.39
58	Asp	8	120.91	177.1	55.02	40.27
59	Lys	7.88	122.47	177.15	57.35	32.87
60	Asn	8.2	118.95	176.01	53.35	39.08
61	Gly	8.04	111.09	174.29	46.19	-
62	Asp	8.3	122.25	176.96	54.24	41.34
63	Gly	8.86	111.5	173.43	45.92	-
64	Glu	7.98	121.21	175.52	55.5	32.25
65	Val	9.01	118.36	175.21	60.67	34.86
66	Ser	8.67	121.66	175.33	57.06	66.61
67	Phe	9.32	124.32	177.26	61.42	38.4
68	Glu	8.52	120.09	179.4	60.33	29.04
69	Glu	7.35	120.63	178.8	59.08	30.25
70	Phe	8.61	123.33	175.71	60.74	39.47
71	Gln	8.03	118.79	178.12	59.93	29.42
72	Val	7.05	120.97	177.96	66.51	31.73
73	Leu	7.74	122.99	178.48	58.6	41.71
74	Val	7.94	117.99	178.38	66.2	31.14
75	Lys	7.43	122.91	178.73	59.63	32.01
76	Lys	7.79	120.29	179.28	59.25	32.73
77	Ile	7.88	116.31	176.52	63.46	38.16
78	Ser	7.61	116.76	174.09	59.57	63.88
79	Gln	7.64	122.24	174.97	56.09	28.96
80	Cys	7.68	125.54	178.59	60.03	29.17

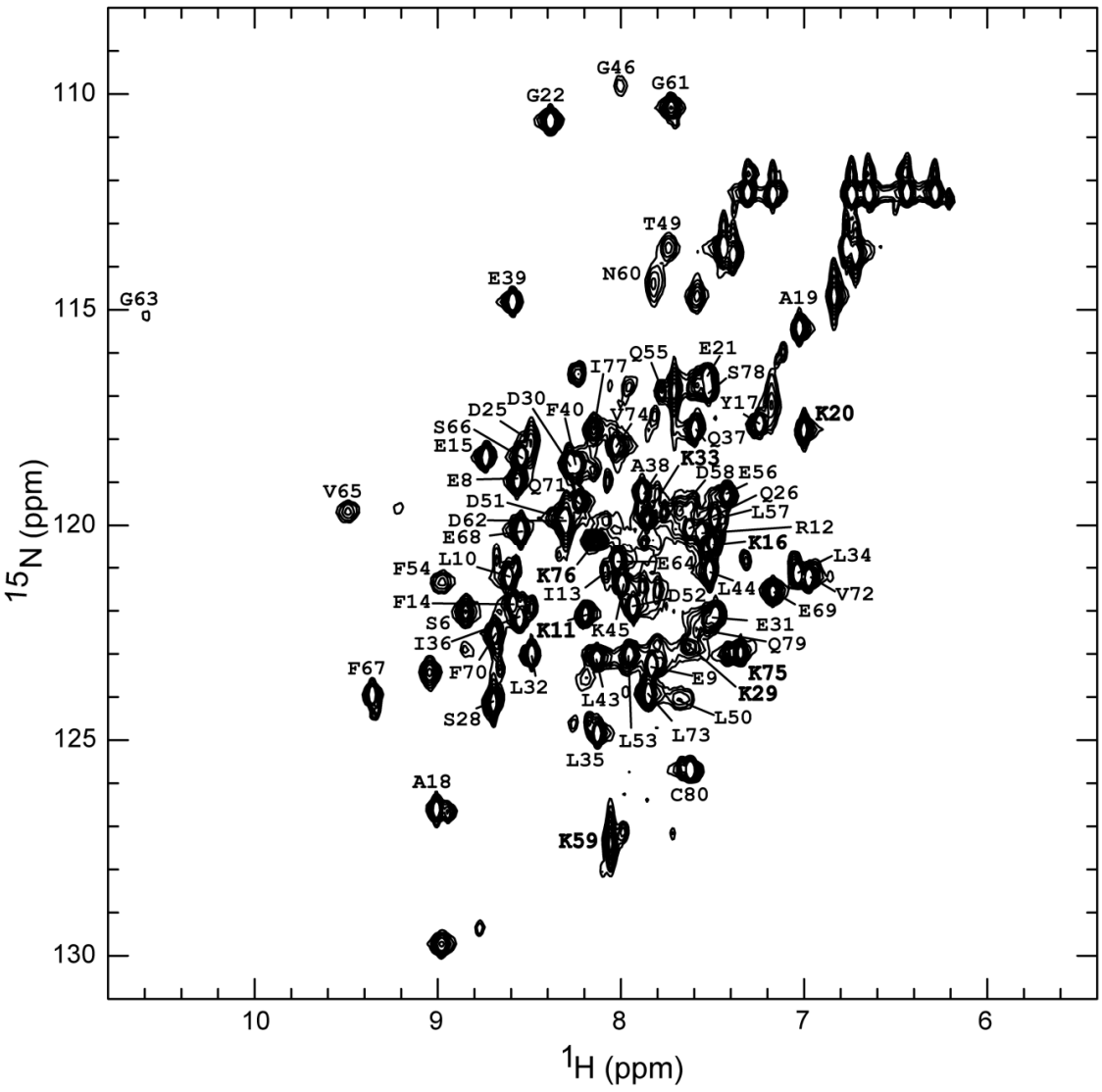


Fig 5.7.4

^{15}N-HSQC showing the backbone resonance assignments of the Mg^{2+} bound human calbindin D_{9k} (P47M+C80).

Table 5.7.2

Backbone resonance assignment of the Mg^{2+}bound human calbindin D_{9k}(P47M+C80) *in vitro* at 37°C

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
1	Met	-	-	-	-	-
2	Ser	-	-	-	-	-
3	Thr	-	-	-	-	-
4	Lys	-	-	-	-	-
5	Lys	8.29	123.83	176.3	55.17	34.33
6	Ser	8.84	121.94	173.05	56.82	63.3
7	Pro	-	-	179.12	66.18	31.92
8	Glu	8.55	118.97	179.1	60.1	29.14
9	Glu	7.83	123.16	179.42	59.09	30.19
10	Leu	8.6	121.09	178.79	58.4	42.1
11	Lys	8.2	122.07	177.67	59.55	31.97
12	Arg	7.54	120.07	179.87	59.42	30.41
13	Ile	7.98	121.22	176.81	65.52	38.82
14	Phe	8.59	121.85	176.83	62.8	40.06
15	Glu	8.7	118.28	178.91	59.59	29.91
16	Lys	7.49	120.33	177.75	59.07	32.34
17	Tyr	7.22	117.48	178.04	60.81	39.77
18	Ala	8.98	126.33	178.23	54.74	19.17
19	Ala	7.02	115.57	178.76	52.97	18.78
20	Lys	7.02	117.94	176.46	59.52	33.02
21	Glu	7.54	116.44	174.92	55.02	32.71
22	Gly	8.35	110.63	173.99	45.94	-
24	Pro	-	-	-	64.08	31.48
25	Asp	8.5	118.12	175.65	53.37	41.94
26	Gln	7.58	120	173.24	-	-

Table 5.7.2 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
27	Leu	-	-	-	-	-
28	Ser	8.68	124.2	174.67	57.58	65.26
29	Lys	7.63	122.79	177.72	60.78	31.49
30	Asp	8.25	118.54	179.36	57.74	40.5
31	Glu	7.47	121.92	178.39	58.68	29.91
32	Leu	8.47	122.83	178.04	57.91	41.04
33	Lys	7.87	119.85	178.56	61.12	32.21
34	Leu	7.02	121	178.57	57.98	41.31
35	Leu	8.08	124.7	179.46	59.22	42.27
36	Ile	8.56	122.11	177.39	65.29	37.15
37	Gln	7.59	117.69	176.83	59.2	29.11
38	Ala	7.85	119.1	179.95	54.25	20.36
39	Glu	8.56	114.7	176.79	55.98	30.5
40	Phe	8.24	118.57	173.14	54.72	40.65
41	Pro	-	-	-	65.38	31.2
42	Ser	8.27	115.58	175.68	60.45	62.68
43	Leu	8.12	122.99	177.61	55.53	41.33
44	Leu	7.51	121.06	178.42	56.91	41.76
45	Lys	7.94	121.53	177.21	57.53	32.22
46	Gly	7.99	109.72	-	45.44	-
47	Met	-	-	-	-	-
48	Asn	-	-	176.26	55.51	37.96
49	Thr	7.73	113.48	175.38	62.52	69.66
50	Leu	7.67	123.99	177.85	58.61	42.12
51	Asp	8.31	119.75	178.74	58.65	40.27
52	Asp	7.92	121.83	178.93	57.35	40.51
53	Leu	7.96	122.98	178.6	57.58	42.83
54	Phe	8.96	121.35	176.68	62.73	39.79

Table 5.7.2 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
55	Gln	7.7	116.73	177.87	58.68	28.63
56	Glu	7.41	119.17	177.76	58.62	29.9
57	Leu	7.47	119.63	177.27	55.51	42.81
58	Asp	7.66	119.53	178.54	52.44	39.49
59	Lys	8.06	127.37	177.69	59.05	32.02
60	Asn	7.81	114.22	176.59	51.89	36.73
61	Gly	7.7	110.21	174.41	47.78	-
62	Asp	8.29	119.72	177.26	52.99	40.8
63	Gly	10.57	115.26	173.99	45.93	-
64	Glu	7.99	120.49	175.43	54.45	33.45
65	Val	9.5	119.84	175.08	59.14	35.62
66	Ser	8.5	118.08	175.4	56.48	66.5
67	Phe	9.34	123.88	177.1	61.26	38.4
68	Glu	8.54	120.13	179.57	60.25	29.07
69	Glu	7.16	121.55	178.14	58.95	29.54
70	Phe	8.69	122.41	175.64	60.71	39.51
71	Gln	8.22	119.4	178.3	59.95	29.43
72	Val	6.96	121.18	177.34	66.48	31.43
73	Leu	7.83	123.9	178.3	59.95	41.84
74	Val	8.02	118.08	178.19	66.42	31.21
75	Lys	7.33	122.89	179.08	59.65	32.02
76	Lys	8.12	120.24	179.76	59.38	32.74
77	Ile	8.13	117.67	176.59	63.78	38.12
78	Ser	7.51	116.82	174.17	59.88	63.83
79	Gln	7.52	122.15	175.03	56.02	28.99
80	Cys	7.61	125.62	178.62	60.22	29.08

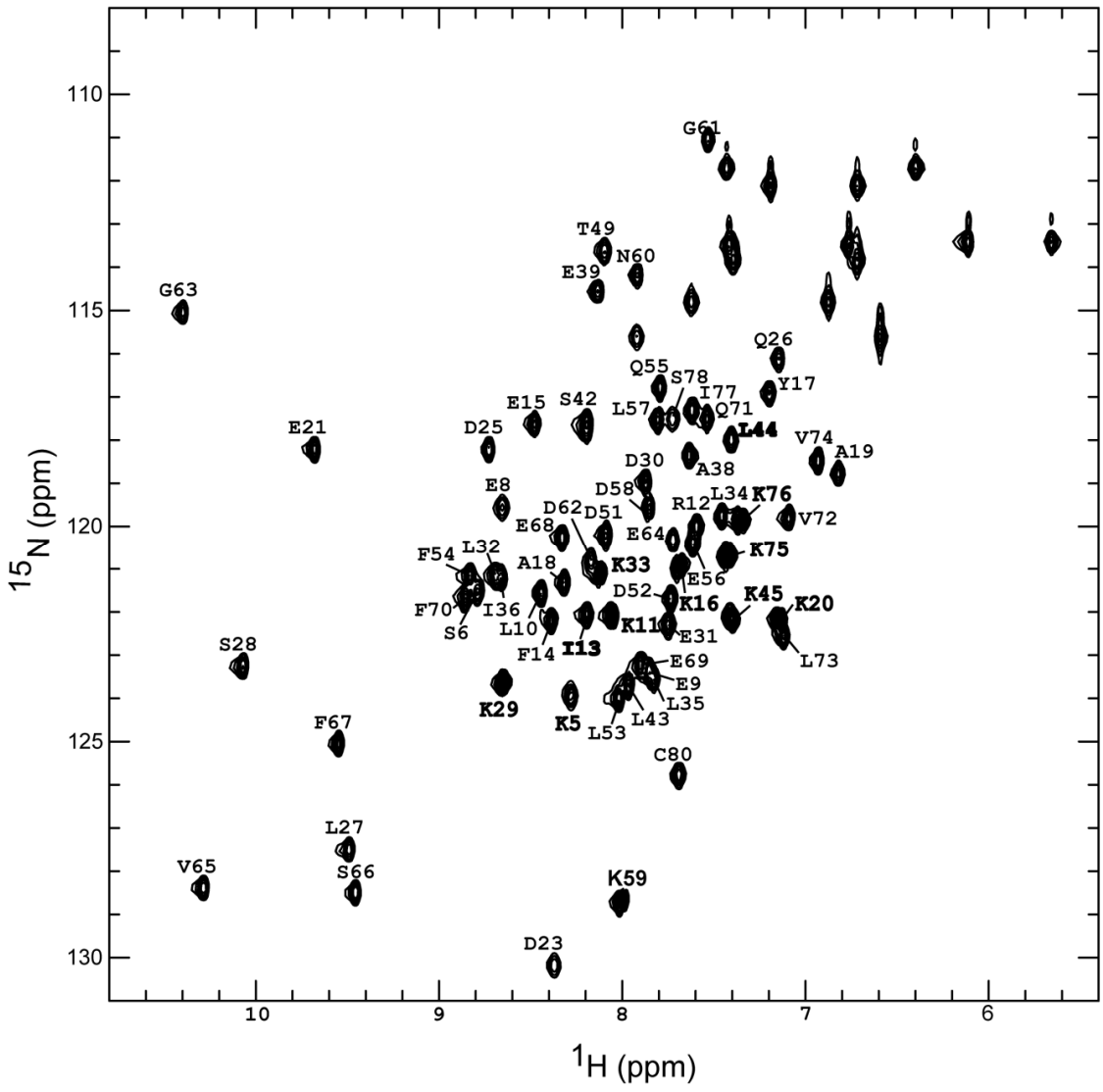


Fig 5.7.5

^{15}N-HSQC showing the backbone resonance assignment of the Ca^{2+} bound human calbindin D_{9k} (P47M+C80).

Table 5.7.3

Backbone resonance assignment of the Ca^{2+}bound human calbindin D_{9k}(P47M+C80) *in vitro* at 37°C

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
1	Met	-	-	-	-	-
2	Ser	-	-	-	-	-
3	Thr	-	-	-	-	-
4	Lys	-	-	176.37	56.56	33.14
5	Lys	8.35	124.13	175.92	55.27	34.55
6	Ser	8.79	121.49	173.22	56.61	63.25
7	Pro	-	-	179.51	66.17	31.92
8	Glu	8.64	119.46	179.16	60.19	29.13
9	Glu	7.85	123.21	179.44	59.13	30.15
10	Leu	8.44	121.41	178.64	58.3	42.14
11	Lys	8.07	121.99	177.59	59.94	32.26
12	Arg	7.59	119.9	179.93	59.51	30.23
13	Ile	8.18	121.94	176.74	65.88	38.75
14	Phe	8.39	122.1	176.54	63.05	39.66
15	Glu	8.48	117.56	179.18	59.14	29.92
16	Lys	7.69	120.91	178.01	58.96	32.31
17	Tyr	7.18	116.84	176.2	61.51	39.22
18	Ala	8.3	121.21	178.98	54.23	17.76
19	Ala	6.83	118.82	177.64	52.6	19
20	Lys	7.16	122.1	177.23	60.33	32.75
21	Glu	9.66	118.12	176.14	54.23	33.21
22	Gly	8.87	114.91	173.99	46.06	-
23	Asp	8.35	129.97	-	52.4	41.32
24	Pro	-	-	174.49	64.23	31.46
25	Asp	8.71	118.13	175.1	53.47	42.64

Table 5.7.3 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
26	Gln	7.14	116.05	174.31	55.03	35.88
27	Leu	9.5	127.49	176.63	53.21	44.44
28	Ser	10.06	123.17	175.33	56.43	65.73
29	Lys	8.65	123.56	178.22	61.27	31.17
30	Asp	7.87	118.86	179.31	57.53	41
31	Glu	7.74	122.22	178.79	59.53	30.69
32	Leu	8.68	121.06	177.72	57.92	41.05
33	Lys	8.12	121.04	177.71	60.71	32.04
34	Leu	7.45	119.68	178.25	57.79	42.06
35	Leu	7.82	123.47	179.04	58.89	42.58
36	Ile	8.66	121.17	177.92	65.5	37.39
37	Gln	8.19	117.56	177.15	59.39	28.83
38	Ala	7.62	118.29	179.76	54.2	20.72
39	Glu	8.13	114.48	176.75	56.22	30.96
40	Phe	8.2	117.76	173.18	54.92	41.44
41	Pro	-	-	-	65.42	31.22
42	Ser	8.31	116.37	-	60.66	62.55
43	Leu	7.96	123.58	177.9	57.1	42.59
44	Leu	7.4	117.95	176.8	54.69	41.24
45	Lys	7.41	122.02	176.73	56.48	32.94
46	Gly	-	-	-	-	-
47	Met	-	-	-	-	-
48	Asn	-	-	175.32	53.98	39.47
49	Thr	8.09	113.6	175.24	61.53	71.22
50	Leu	-	-	178.36	57.34	41.25
51	Asp	8.08	120.11	178.54	58.26	40.54
52	Asp	7.73	121.59	178.86	57.39	40.54
53	Leu	8.01	123.9	178.16	57.65	42.08

Table 5.7.3 *continued*

Residue No.	Type	$^{1}H^{N}$	^{15}N	$^{13}C'$	$^{13}C^{\alpha}$	$^{13}C^{\beta}$
54	Phe	8.82	121.08	176.29	63.47	39.73
55	Gln	7.79	116.71	177.89	58.66	28.63
56	Glu	7.62	120.34	178.1	58.66	29.62
57	Leu	7.8	117.54	179.16	55.79	43.42
58	Asp	7.84	119.49	177.09	52.98	38.21
59	Lys	8.01	128.68	178.12	58.42	33.09
60	Asn	7.92	114.15	176.35	51.93	37.03
61	Gly	7.53	110.96	174.84	47.77	-
62	Asp	8.16	120.7	177.62	53.22	40.54
63	Gly	10.39	114.99	172.83	46.02	-
64	Glu	7.73	120.21	175.97	54.01	34.71
65	Val	10.26	128.3	175.62	61.4	33.43
66	Ser	9.45	128.41	175.12	57.36	66.28
67	Phe	9.54	124.98	177.66	61.98	38.17
68	Glu	8.33	120.11	180.37	59.92	29.42
69	Glu	7.88	123.08	179.53	58.92	29.88
70	Phe	8.84	121.63	175.96	60.18	40.23
71	Gln	7.54	117.38	177.53	58.09	27.87
72	Val	7.08	119.68	177.86	65.82	31.98
73	Leu	7.12	122.4	177.8	57.17	41.34
74	Val	6.93	118.37	177.8	66.27	31.29
75	Lys	7.43	120.62	178.32	58.95	32.23
76	Lys	7.37	119.7	178.68	58.64	33.02
77	Ile	7.61	117.17	176.91	62.92	38.3
78	Ser	7.74	117.64	-	59.64	63.83
79	Gln	7.72	122.23	174.99	55.95	29.32
80	Cys	7.69	125.62	178.6	60	29.2

5.8 Lysine-selectively ^{15}N-labelling of human calbindin D_{9k}(P47M+C80).

Lysine selectively ^{15}N-labelling of human calbindin $D_{9k}$was performed for the reduction of ambiguity in the assignment transfer process for the in-cell NMR experiments (chapter 7.3). As there are 11 lysine residues in the human calbindin D_{9k}(P47M+C80), it became easy to transfer the assignment from the uniformly ^{15}N-labelled "assigned" spectra to the lysine-selectively labelled spectra. Selectively labelled spectra were assigned for the metal free, Mg^{2+} bound, and Ca^{2+} bound of human calbindin D_{9k}(P47M+C80). The spectra of uniformly labelled and selectively labelled calbindin D_{9k}(P47M+C80) in the metal free state were properly superimposed in order to confirmed their positions for metal free (fig 5.8.1). Likewise, the assignments were successfully transferred to the 2D ^{1}H-^{15}N HSQC spectra of lysine-selectively ^{15}N-labelled samples in the Mg^{2+} bound (Fig 5.8.2) and Ca^{2+}bound (Fig 5.8.3) forms, respectively.

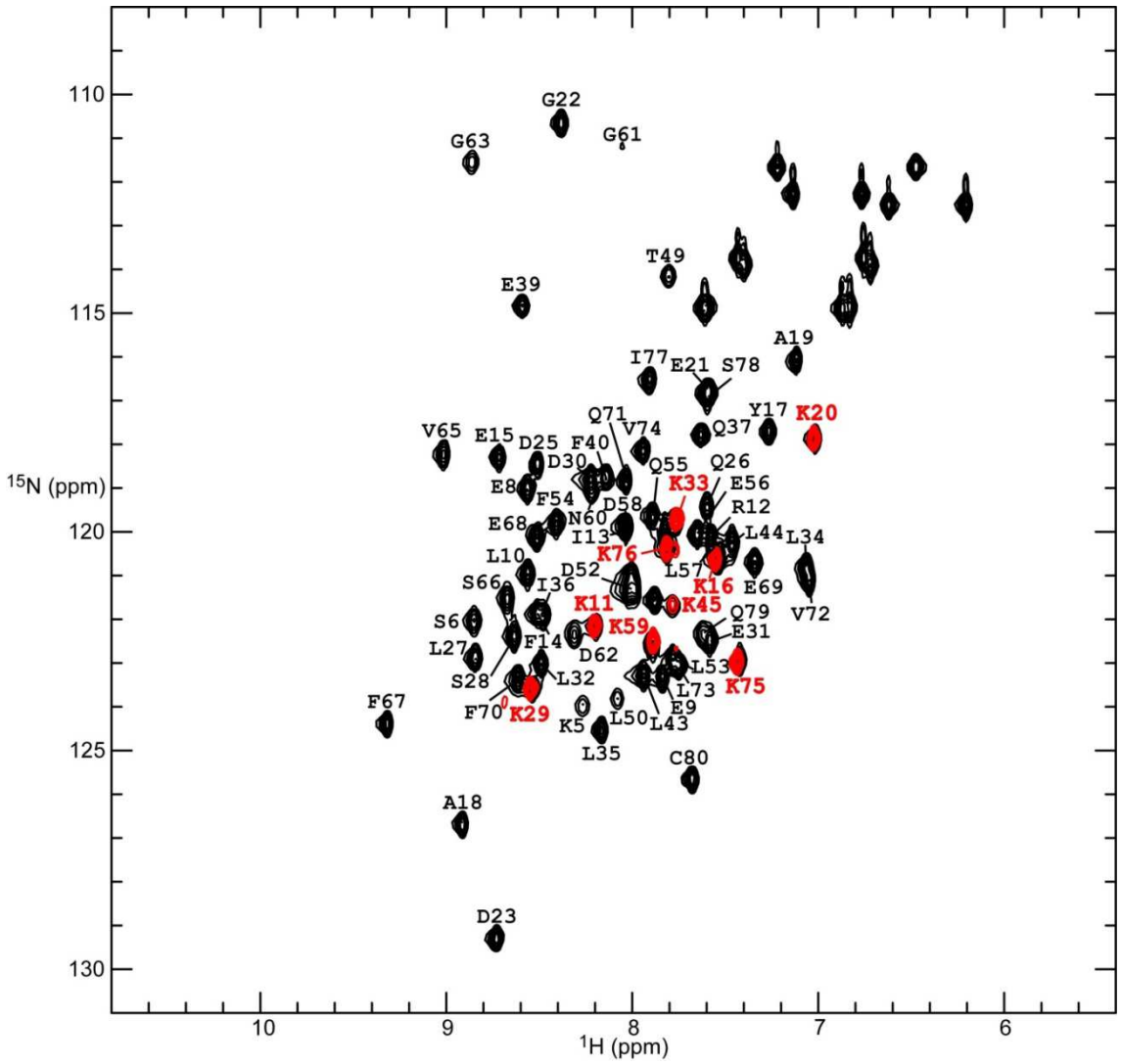


Fig. 5.8.1

Superimposed ^{15}N-HSQC spectra of uniformly labelled (black) and lysine-selectively labelled (red) calbindin D_{9k} (P47M+C80) in the metal-free state.

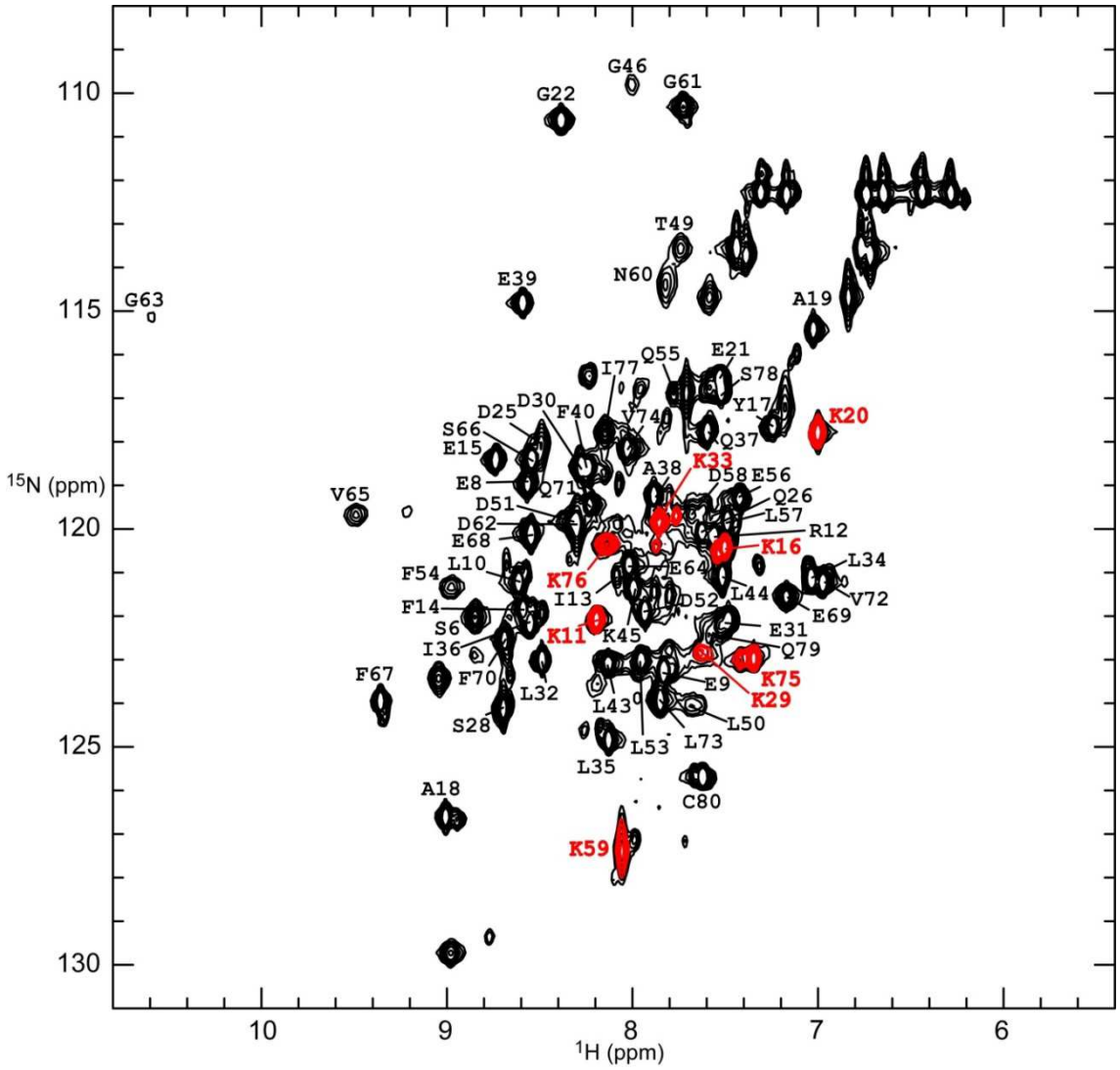


Fig. 5.8.2

Superimposed ^{15}N-HSQC spectra of uniformly labelled (black) and lysine-selectively labelled (red) human calbindin D_{9k} (P47M+C80) in the Mg^{2+}-bound state.

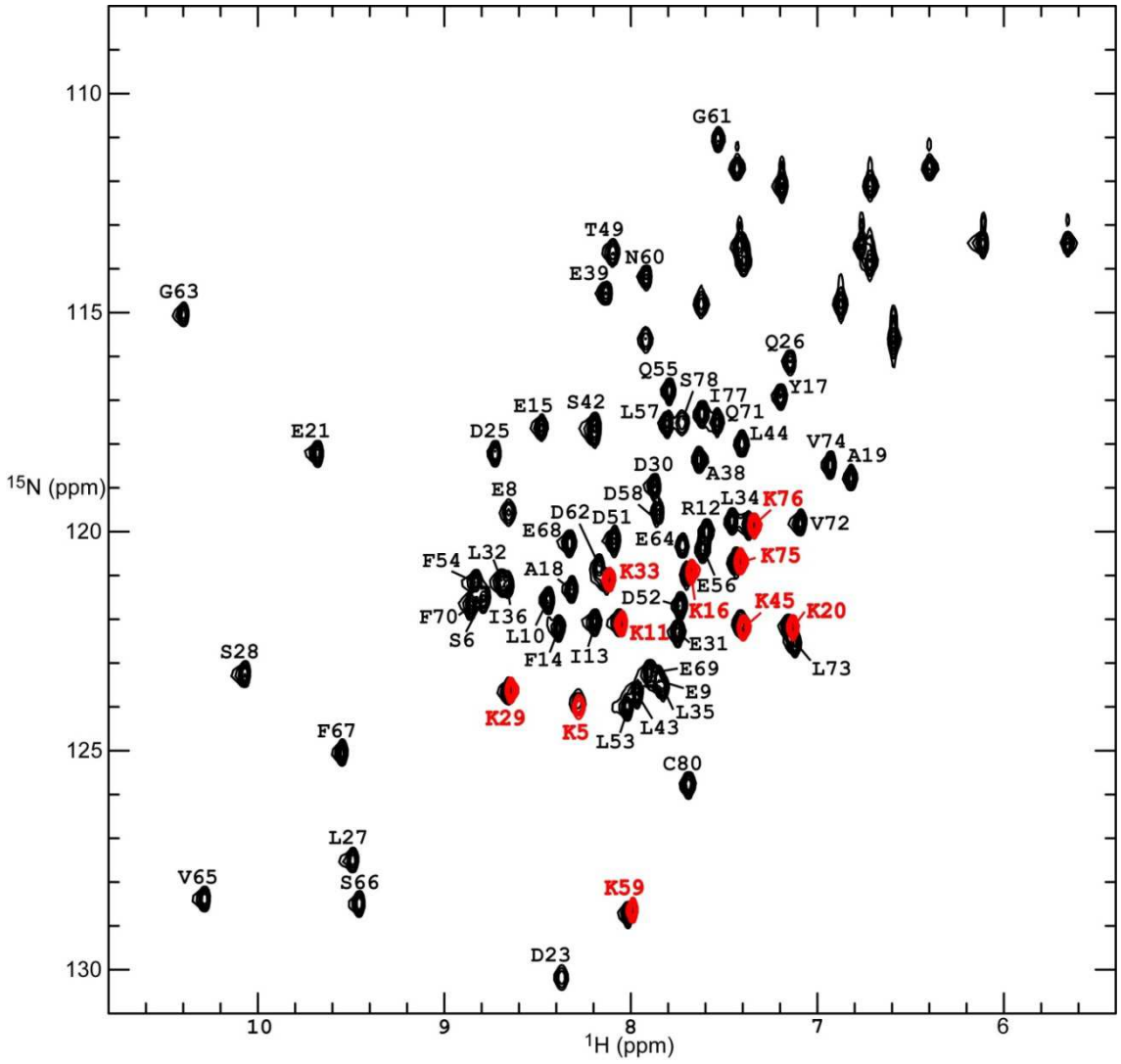


Fig. 5.8.3

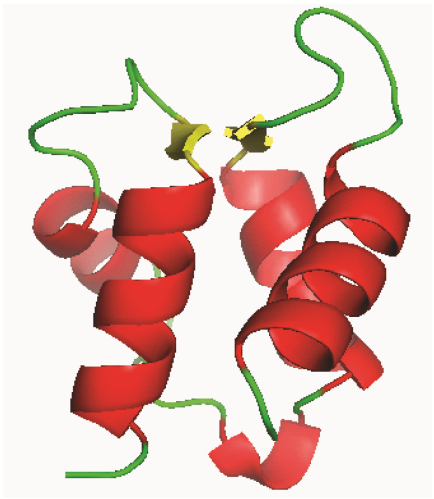
Superimposed ^{15}N-HSQC spectra of uniformly labelled (black) and lysine-selectively labelled (red) human calbindin D_{9k} (P47M+C80) in the Ca^{2+}-bound state.

5.9 Homology modelling

Homology modelling is a comparative modelling of protein, to construct an atomic-resolution model of the "*target*" protein from its amino acid sequence by using an experimentally determined three-dimensional structure of a related homologous protein as a template protein. Homology modelling relies on the identification of one or more known protein structures likely to resemble the structure of the query sequence, and on the production of an alignment of sequence that maps residues in the query sequence to residues in the template sequence. Evolutionarily related proteins have similar sequences and naturally occurring homologous proteins have similar protein structures. It has been shown that three-dimensional protein structures are evolutionarily more conserved than would be expected on the basis of sequence conservation alone [79], while the sequences falling below a 20% sequence identity can have very different structure [80]. Homology models can also be used to identify subtle differences between related proteins that have not all been solved structurally.

As the human calbindin D_{9k} structure was not solved yet, I used a program called “MODELLER” version 9.12 to generate the structure of human calbindin D_{9k}, using bovine calbindin D_{9k} (PDBID: 1CLB) as template.

From the backbone assignment, the chemical shift perturbation for the backbone $^{1}H^{N}$ and ^{15}N nuclei of metal free calbindin D_{9k}(P47M+C80) upon Ca^{2+} and Mg^{2+} binding were calculated and shown in Fig. 5.9.2.



Bovine calbindin D_{9k} used as template for homology modeling. PDBID: 1CLB

Human calbindin D_{9k} (P47M+C80) generated by homology modeling.

Fig 5.9.1

The structures was generated by homology modelling on a template structure of metal free Bovine calbindin (PDBID:1CLB) with the program MODELLER version 9.12.

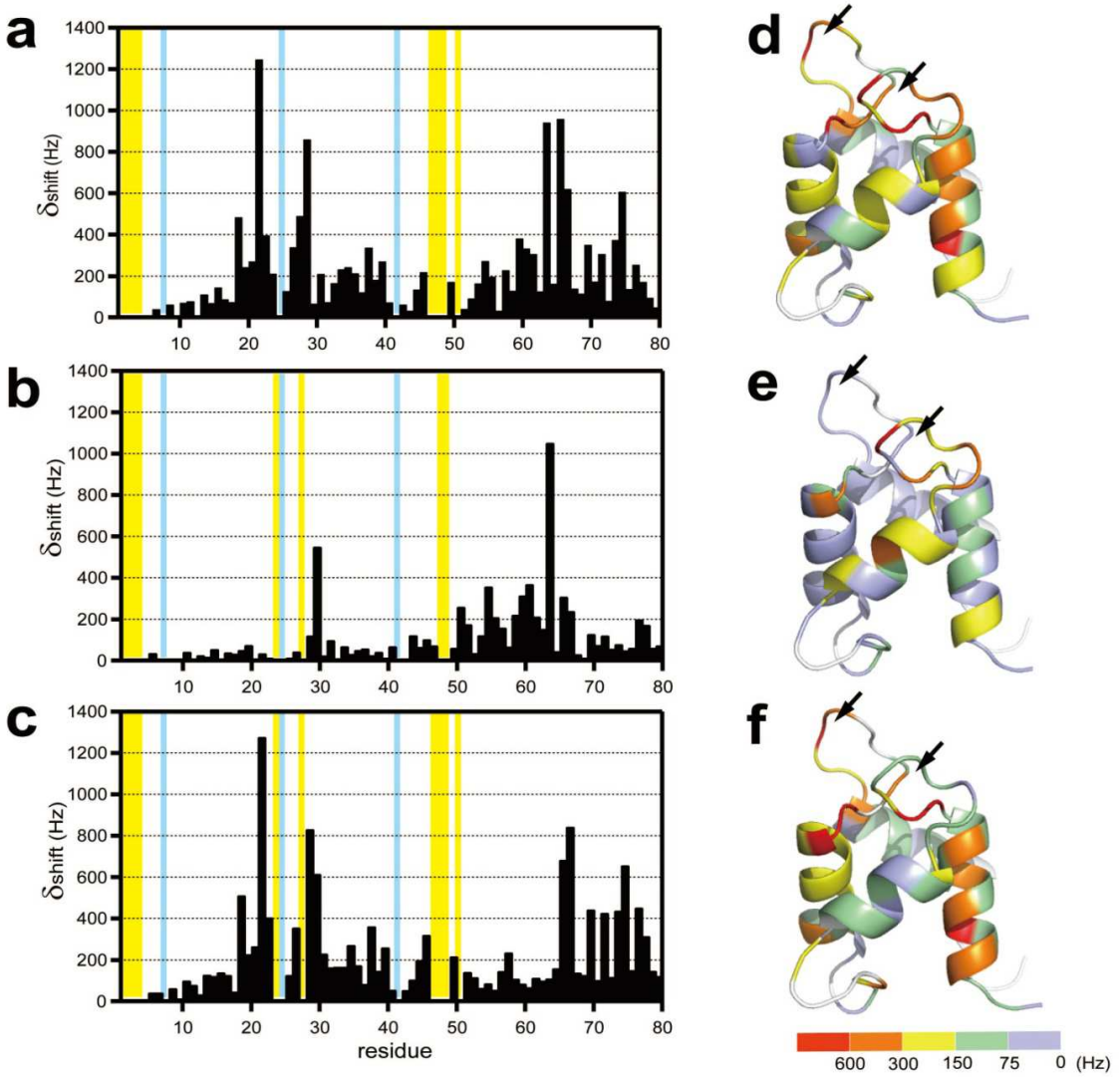


Fig. 5.9.2.

Chemical shift perturbation for backbone $^{1}H^{N}$ and ^{15}N nuclei of metal free calbindin D_{9k}(P47M+C80) upon the calcium (**a,d**) and magnesium bound form (**b,e**), and of calcium bound calbindin D_{9k}(P47M+C80) upon the magnesium bound form (**c,f**). The mean shift difference, δ_{shift}, for each amino acid residue was calculated as $[(\delta^{1}H^{N})^{2}+(\delta^{15}N)^{2}]^{1/2}$, where $\delta^{1}H^{N}$ and$\delta^{15}N$ are the chemical shift differences (Hz) for H^{N} and ^{15}N resonances observed at 14.1T (^{1}H frequency is 600 MHz). The colour coding on the ribbon diagrams is as follows: white (no assignment); light blue ($0<\delta_{shift}\leq75$ Hz); light green ($75<\delta_{shift}\leq150$ Hz); yellow ($150<\delta_{shift}\leq300$ Hz); orange ($300<\delta_{shift}\leq600$ Hz); red ($\delta_{shift}>600$ Hz).

The arrows show the metal binding sites. The structures were calculated by homology modelling on a template structure of metal free Bovine calbindin (PDBID:1CLB) with the program MODELLER version 9.12. Molecular graphics images were produced using the program PYMOL version 1.3.

Chapter 6

In-cell NMR Study in HeLa Cells

6.1 Sample preparation for in-cell NMR experiments.

As discussed in the chapter 3, for in-cell NMR methods for HeLa cells, human calbindin D_{9k} was prepared (discussed in chapter 5). The buffer of purified human calbindin D_{9k} was changed to PBS (phosphate buffer saline), confirming no reducing agent is present in the buffer. The CPP sequence of Tat protein from HIV-1 (CPP-Tat: 47-YGRKKRRQRRR-57) having an additional cysteine at its N-terminal was conjugated to the cysteine residue at carboxyl terminus of human calbindin D_{9k}(P47M+80C), followed by a purification by using PD10 column (GE Healthcare). SDS-PAGE (loading buffer must be free from reducing agent like DTT) confirms the conjugation of human calbindin D_{9k} with CPP-TAT. The purified human calbindin D_{9k}-Cys-Cys-CPP-TAT was incubated (approximately 10 minutes) with HeLa cells (sub-cultures on a 9 cm diameter petri dish, and are sticky to the plate) in the presence of

pyrenbutyrate (1-pyrenebutyric acid), which mediates the direct translocation of CPP-linked human calbindin D_{9k} into the cytosol of the HeLa cells. This incubation process was repeated four times with an interval of 40 minutes. The incubation process was done in a humidified CO_2 incubator at $37^{O}C$ in Dulbecco's Modified Eagle Medium (DMEM) buffer with 10% fatal bovine serum (FBS) and 1% penicillin-streptomycin, hence abbreviated as DMEM(+). Waiting for approx. 40 minutes after the final incubation, the HeLa cells were treated with Trypsin/EDTA in order to remove the cells from the dish plate. The cells were transferred to a 50 mL falcon tube containing freshly prepared medium containing DMEM(+) with HEPES (5mM) and D-glucose (90 mM), and then centrifuged at 2 G for 5 minutes in room temperature and this process was repeated once more. Removing the supernatant by aspiration, the cells were re-suspended with 100 μL of final in-cell NMR buffer which contains DMEM buffer, 5 mM of HEPES, 90 mM D-glucose and 10% D_2O, and transferred to a 3mm Shigemi tube. The falcon tube was washed with 100 μL of NMR buffer again to collect the remaining cells and transferred to the same NMR tube. This loading process of cells onto the NMR tube leads to two different zone where the lower part of the NMR tube will be having highly densed cells and the upper part will be less. During the NMR measurement time all the cells will be precipitated at the bottom of the NMR tube and only buffer will be above the cells. This process doesn't affect the NMR measurement but the physiological property of the protein may change and could lead to observe different chemical shift on the NMR spectra.

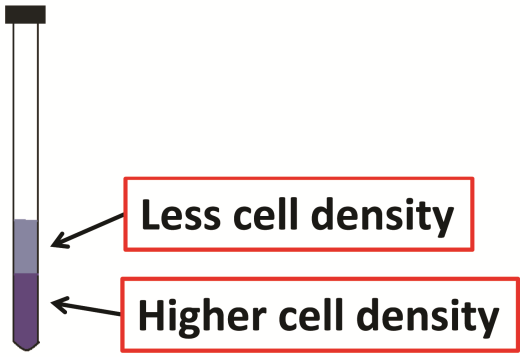


Fig.6.1.1

A schematic drawing of an in-cell NMR sample in an NMR tube after loading the cells creates two zones. The lower part is higher cell density than the upper part.

6.2 In-cell NMR measurement

We performed the in-cell experiments by measuring a series of 2D ^{1}H-^{15}N-SOFAST (band-Selective Optimized-Flip Angle Short Transient) HMQC spectra, which were measured in combination with a non-linearly sampling scheme in its indirect ^{15}N-dimension.

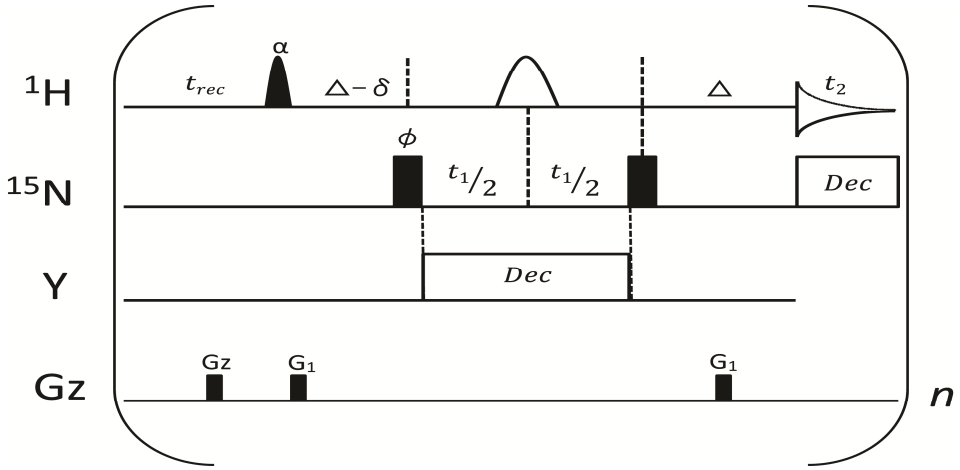


Fig. 6.2.1

^{1}H-^{15}N-SOFAST-HMQC NMR experiment. Filled and open pulse symbols indicate 90° and 180° rf pulses, except for the first ^{1}H excitation pulse applied with flip angleα. The variable flip-angle has a polychromatic PC9 shape and band-selective ^{1}H re-focusing pulse is by REBURP [81]. Quadrature detection in t_1 is obtained by phase incrementation of ϕaccording to TPPI-STATES. Adiabatic WURST-2 decoupling is applied on ^{15}N during detection.

These methods were used to acquire the spectra in a short time due to short living time of HeLa cells in NMR tube. All the NMR measurements were performed at 37°C on a Bruker AVANCE III 600 MHz NMR spectrometer equipped with a cryogenic H/C/N probehead. The spectra were processed with MaxEnt method by Azara software version 2.8.

Chapter 7

Results and discussions

7.1 ^{1}H-^{15}N SOFAST-HMQC spectra of human calbindin D_{9k}(P47M+C80) in HeLa cells.

^{15}N-labelled calbindin D_{9k}(P47M+C80) in the Ca^{2+}-bound state was efficiently incorporated into HeLa cells with the help of C-terminal conjugated CPP^{TAT}. Fig.7.1.1 shows a schematic diagram of the in-cell NMR experiments of calbindin D_{9k}(P47M+C80) in HeLa cells. The concentration of calbindin D_{9k}(P47M+C80) in an in-cell NMR sample was roughly estimated to be ~100 μM. Fig 7.1.2-a show a 2D ^{1}H-^{15}N SOFAST-HMQC spectrum of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells. The ^{1}H-^{15}N correlation cross peaks were well-resolved, suggesting that the proteins in side HeLa cells are properly folded.

7.1.1 Confirmation of in-cell sample, protein leakage and cell viability

It is necessary to be confirmed that the proteins observed in the spectra were indeed inside the living cells. Hence after the in-cell NMR measurements the cells were taken out and centrifuged at 1000 rpm. The supernatant was removed and the precipitated cells were preserved for further investigations. The supernatant was centrifuged at 13000 rpm for complete precipitation of remained cells, and the soluble fraction was used for NMR measurements. Almost all ^{1}H-^{15}N correlation cross peaks disappeared in the spectrum recorded from supernatant (fig 7.1.2–b), confirming that the NMR signals were indeed from the proteins inside the cells. The preserved cells were lysed and the lysate was check by measuring NMR spectra. The essentially identical spectrum (Fig.7.1.2-c), with the exception of shaper line shape, was observed. These results demonstrated that the contribution of extracellular proteins to the observed signals is negligible. The viability of the HeLa cells after approximately 3.5 hours of NMR measurements was confirmed to be 90±3 % by trypan-blue staining.

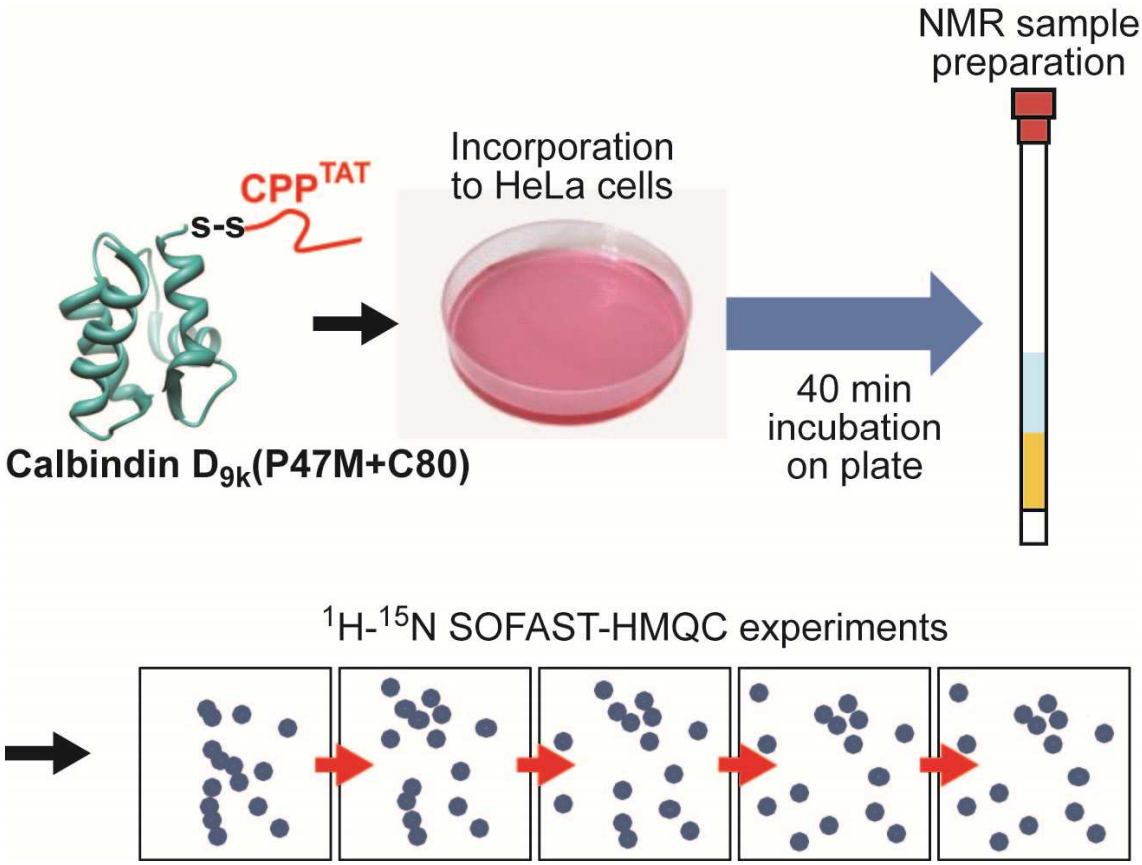


Fig 7.1.1

Schematic illustration of the in-cell NMR experiments of calbindin D_{9k}(P47M+C80) in HeLa cells.

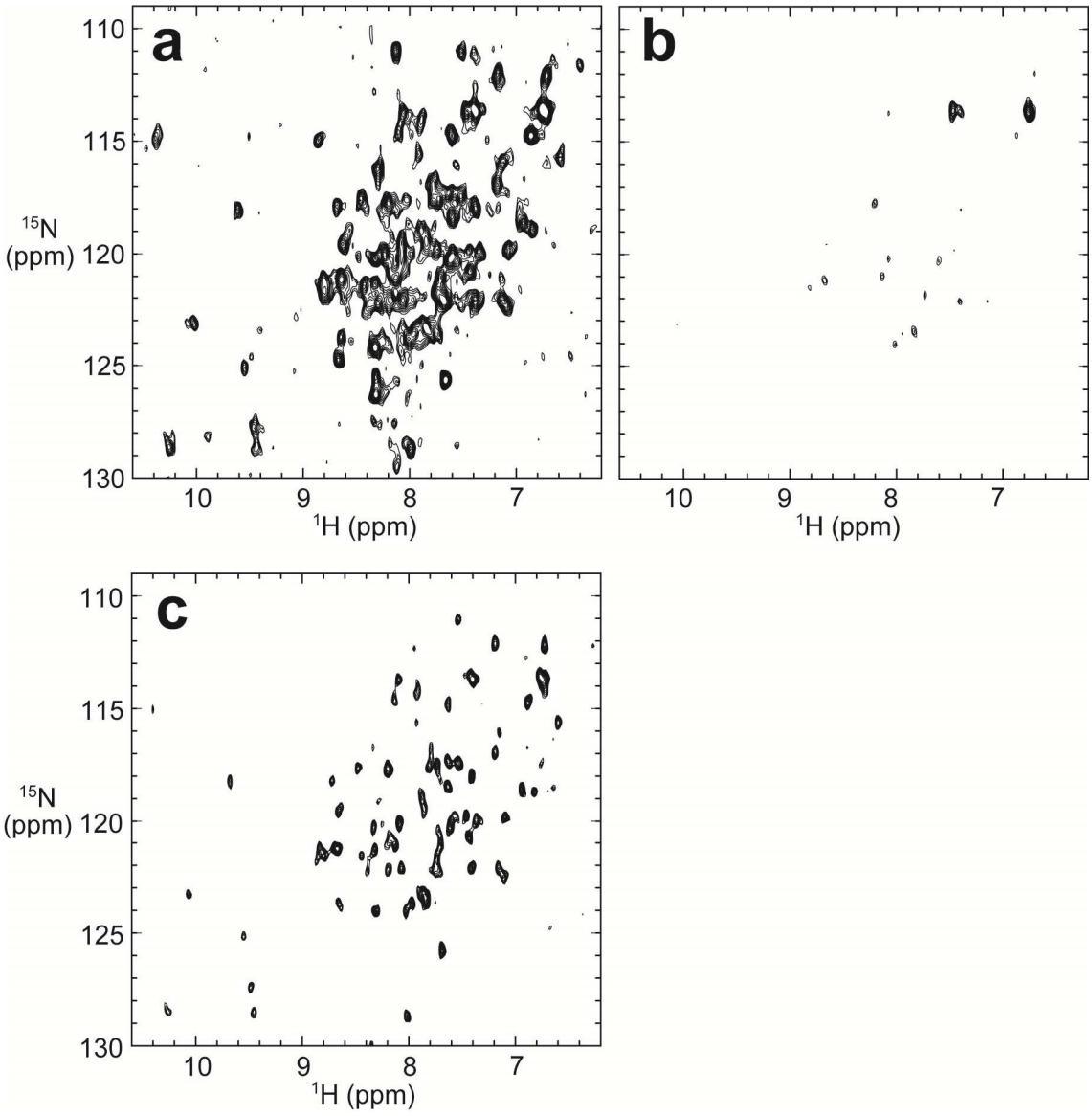


Fig 7.1.2

(**a**) a well spread 2D ^{1}H-^{15}N HSQC spectrum of human calbindin D_{9k}(P47M+C80) inside HeLa cells, (**b**) a spectrum of the supernatant after the in-cell NMR experiments, (**c**) a spectrum of the lysate of the cells.

7.2 Resonance assignments of ^{1}H-^{15}N correlation cross peaks of human calbindin D_{9k}(P47M+C80) in HeLa cells.

When comparing the sequential spectra (each took 38 minutes) in time-dependent changes in the cross peak pattern were apparently noticed (Fig.7.2.1). We repeated the in-cell NMR experiments employing the identical experimental parameters, and found that these changes were highly reproducible.

In order to investigate these changes in detail, the ^{1}H-^{15}N correlation cross peaks in each in-cell NMR spectrum were analysed in reference to the *in vitro* assignments of calbindin D_{9k}(P47M+C80) in the metal free, Mg^{2+}- and Ca^{2+}-bound states. However, considering the overlap of cross peaks in the in-cell NMR spectra, which is mainly caused by the line broadening due to increased viscosity in the cytosol and specific/nonspecific interactions with various intracellular factors, we anticipated difficulties in the assignment transfer process from the *in vitro* to the in-cell NMR data: solid assignments for isolated peaks, whilst highly ambiguous assignments for overlapped or significantly broadened peaks.

For the assignment transfer process, in-cell spectra were corroborated with *in vitro* resonance assignments for human calbindin D_{9k}(P47M+C80), (See chapter)

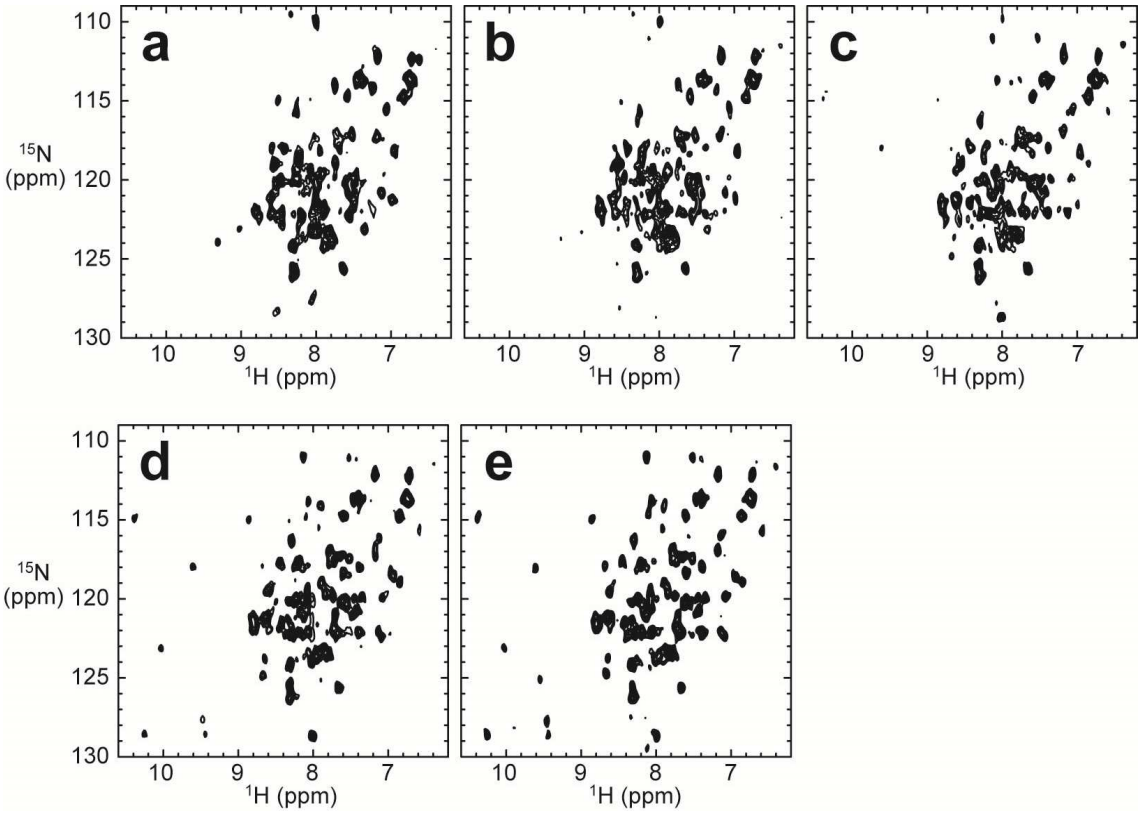


Fig 7.2.1

Time course of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells. The spectra were obtained during the period of 0 to 38 minutes (**a**), 39 to 76 minutes (**b**), 77 to 114 minutes (**c**), 115 to 152 minutes (**d**), and 153 to 190 minutes (**e**) from the start of the in-cell NMR experiments.

7.3 Lysine-selectively ^{15}N-labelling of human calbindin D_{9k} (P47M+C80) and in-cell NMR measurements

As it became difficult to assign and for clear understanding in in-cell NMR measurement, it became necessary to reduce the cross peak in the spectra by selective labelling of the human calbindin D_{9k} and followed by in-cell NMR measurement.

Lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) was prepared and incorporated into HeLa cells for in-cell NMR experiments, for simplifying the spectra, and further assignment transfer process, thus enabling to provide more accurate assignments. Fig.7.3.1 Shows a series of 2D ^{1}H-^{15}N SOFAST-HMQC spectra of lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells. All the NMR experiment parameters are kept same as the experiments on the uniformly ^{15}N-labelled in-cell NMR samples in order to avoid ambiguity.

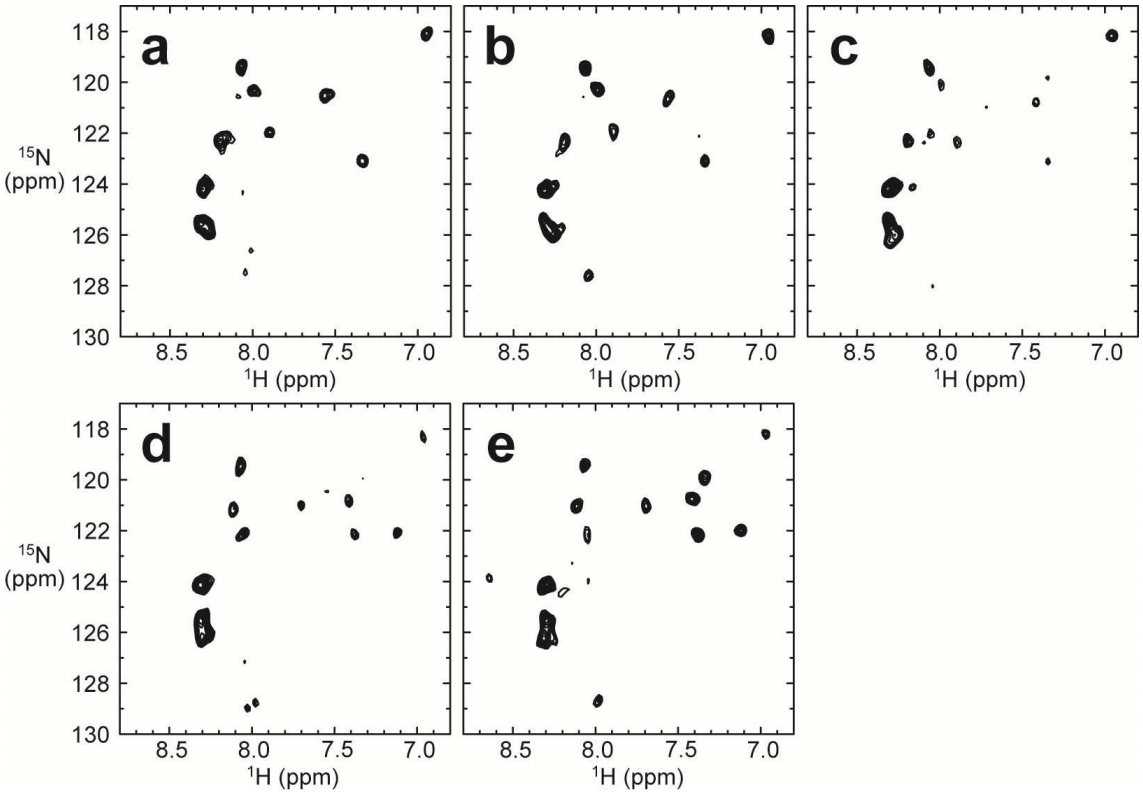


Fig. 7.3.1

Time course of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells. The spectra were obtained during the period of 0 to 38 minutes (**a**), 39 to 76 minutes (**b**), 77 to 114 minutes (**c**), 115 to 152 minutes (**d**), and 153 to 190 minutes (**e**) from the start of the in-cell NMR experiments.

7.4 Time-dependent conversion of calbindin D_{9k}(P47M+C80) from the Mg^{2+}- to the Ca^{2+}-bound states inside HeLa cells.

Fig 7.2.1 shows the time course of the ^{1}H-^{15}N SOFAST-HMQC spectra of the uniformly ^{15}N-labelled human calbindin D_{9k} (P47M+C80) inside HeLa cells. Fig 7.3.1 shows the similar time course of the ^{1}H-^{15}N SOFAST-HMQC spectra of the lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) inside HeLa cells. As calbindin D_{9k} binds various metals it can be hypothesised that this type of change in spectra could be related to its metal- binding (or binding to some other intracellular components). Therefore, the spectra were first compared with the ^{1}H-^{15}N HSQC spectrum of the sample in the Ca^{2+}-bound state *in vitro*.

From the fig: 7.4.1 it is inferred that in the initial stage of the in-cell NMR measurement (0-38 min), the number of observed ^{1}H-^{15}N correlation cross peaks were less than expected, but the cross peak pattern has good agreement with that of the Mg^{2+}-bound state in *in vitro*.

In contrast, the cross peaks observed on uniformly and lysine-selectively [Fig.7.4.2 (a) and (b) respectively] ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells in the later stages (114-152 min from the start of the experiments) showed almost identical chemical shifts with the corresponding cross peaks of the *in vitro* Ca^{2+}-bound calbindin D_{9k}(P47M+C80).

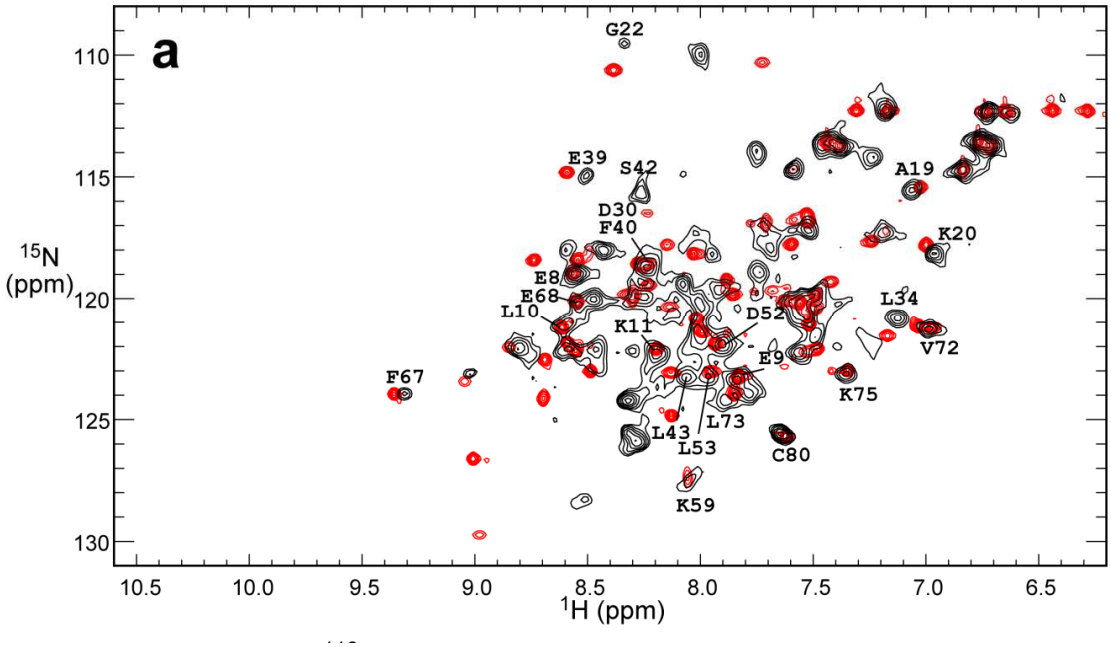


Fig 7.4.1

Overlay of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of calbindin D_{9k}(P47M+C80) in HeLa cells (black) and the 2D ^{1}H-^{15}N HSQC spectra of calbindin D_{9k}(P47M+C80) *in vitro* (red). In the panel **a**, the spectrum of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells which were obtained during the periods of 0 to 38 minutes from the start of the in-cell NMR experiments was compared with uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in the Mg^{2+}-bound state *in vitro*. Likewise, in panel **b**, the spectrum of lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells which were obtained during the periods of 0 to 38 minutes from the start of the in-cell NMR experiments was compared with lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in the Mg^{2+}-bound states *in vitro*. For each panel, cross-peaks in the in-cell NMR spectra are labelled with their corresponding backbone assignments.

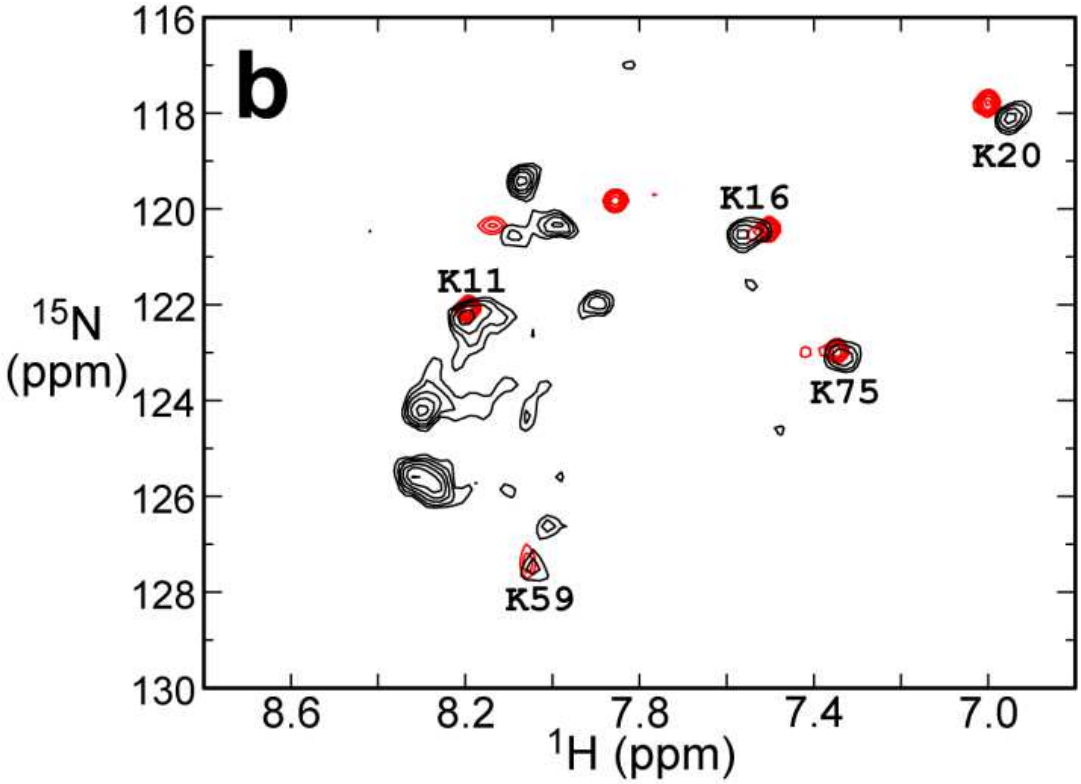


Fig 7.4.1 *continued*

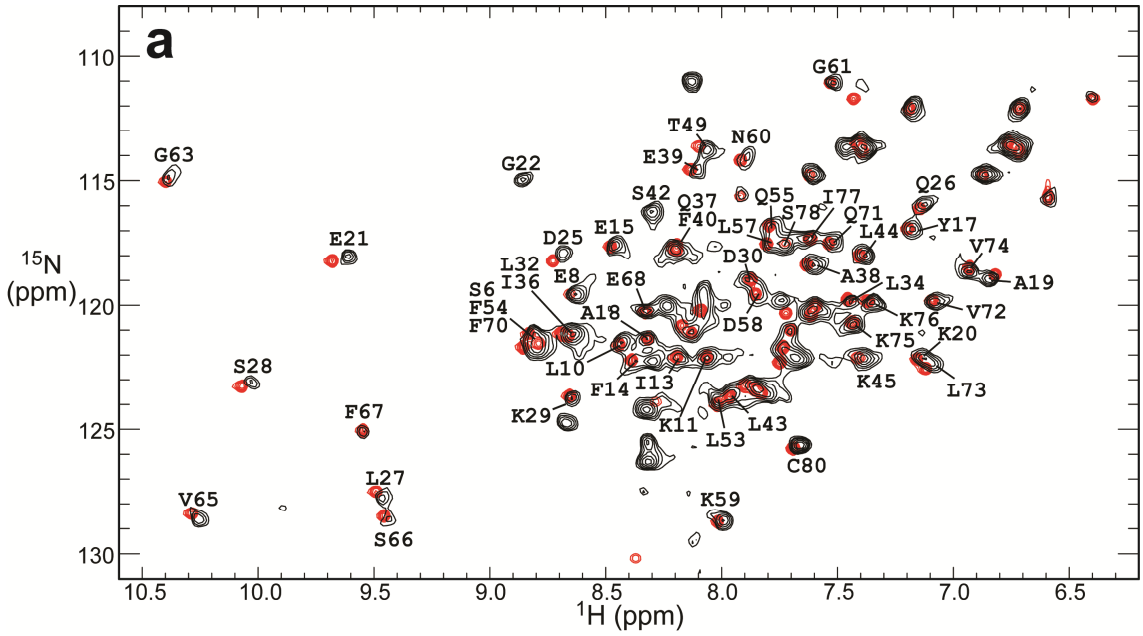


Fig 7.4.2

Overlay of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of calbindin D_{9k}(P47M+C80) in HeLa cells (black) and the 2D ^{1}H-^{15}N HSQC spectra of calbindin D_{9k}(P47M+C80) *in vitro* (red). In the panel **a**, the spectrum of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells which were obtained during the periods of 114-152 minutes from the start of the in-cell NMR experiments was compared with uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in the Ca^{2+}-bound state *in vitro*. Likewise, in panel **b**, the spectrum of lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells which were obtained during the periods of 114-152 minutes from the start of the in-cell NMR experiments was compared with lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) in the Ca^{2+}-bound states *in vitro*. For each panel, cross-peaks in the in-cell NMR spectra are labelled with their corresponding backbone assignments.

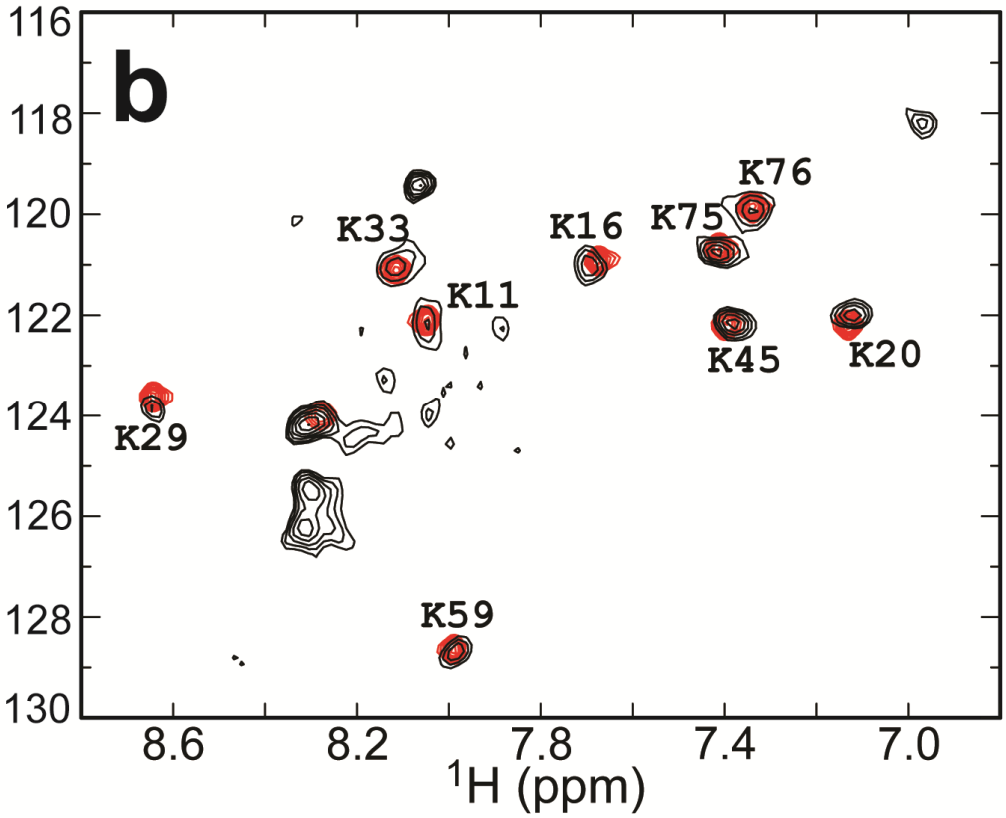


Fig 7.4.2 *continued*

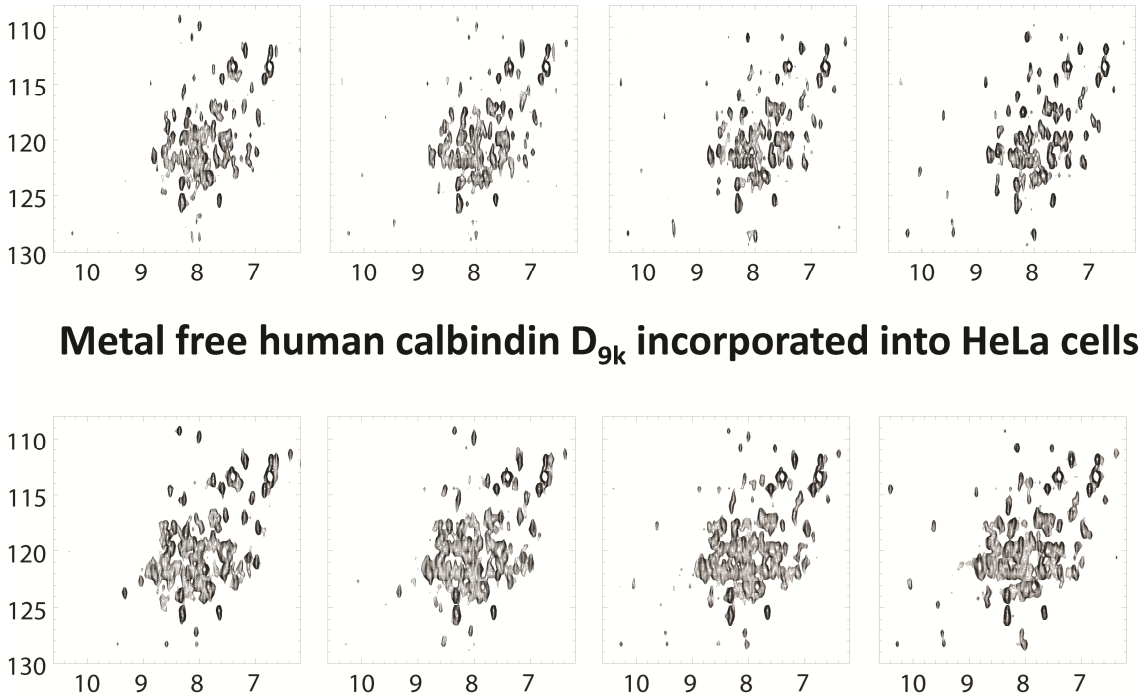


Fig.7.4.3

Time course of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells. In upper panel, the metal-free proteins were used for the incorporation process into HeLa cells. In lower panel, corresponding spectra for which Ca^{2+}-bound calbindin D_{9k}(P47M+C80) was used were shown for comparison. The spectra were obtained during the period of 0 to 38 minutes, 39 to 76 minutes, 77 to 114 minutes and 115 to 152 minutes, respectively from the start of the in-cell NMR experiments.

First, these results demonstrated that the Ca^{2+} ions which were initially bound to the CPP^{TAT}-conjugated calbindin D_{9k}(P47M+C80) were somehow released during the incorporation process into HeLa cells. This was confirmed by the similar in-cell NMR experiments employing the metal-free calbindin D_{9k}(P47M+C80) for the incorporation process, in which identical results were obtained (Fig.7.4.3).

Further our results showed that the calbindin D_{9k}(P47M+C80) just after the incorporation process binds Mg^{2+} ions. Time-dependence in signal intensity of representative cross-peaks suggested that the conversion from Mg^{2+}- to Ca^{2+}-bound states occurred continuously (Fig. 7.4.4).

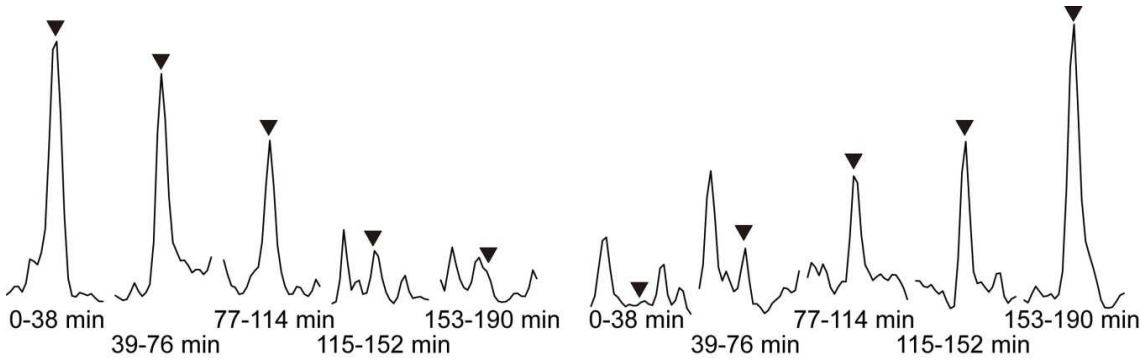


Fig.7.4.4

1D cross sections of 2D ^{1}H-^{15}N SOFAST-HMQC spectra of lysine-selectively ^{15}N-labelled calbindin D_{9k}(P47M+C80) taken at the positions corresponding to the K75 residue in the Mg^{2+}- (left panel) and Ca^{2+}-bound state (right panel), indicating the time-dependent change in cross-peak intensity. Inverted triangles indicate the positions of the amide proton resonances due to the K75 residue.

To interpret the results, we hypothesised that the Ca^{2+}-burst , presumably from Endoplasmic Reticulum (ER) happened continuously after the start of the in-cell NMR experiments, consequently the bound Mg^{2+} ions were replaced by Ca^{2+} ions, since the affinity of calbindin D_{9k} to Ca^{2+} is approximately 10^3-fold higher than Mg^{2+}. In order to investigate when this Ca^{2+}-burst and the resulting time-dependent changes in calbindin D_{9k}(P47M+C80) start during the experiments, we introduced extra incubation time (3 hours) in DMEM(+) on the culture dish between the protein incorporation and the NMR sample preparation processes for in-cell measurement in the protocol (Fig. 7.4.5).

The additional incubation time caused no significant difference in in-cell NMR spectra (Fig7.4.6), demonstrating that the time-dependent changes in the spectra do not start after the incorporation of calbindin D_{9k}(P47M+C80), but initiated after the NMR sample preparation process, in which HeLa cells were removed from the culture dish by trypsin/EDTA, washed and placed into an NMR tube.

As was mentioned above Ca^{2+}-burst occurs in HeLa cells under various stresses. It is known that NMR radio frequency pulses, which continuously move and reorient ions and polar molecules by the alternating electric field, thereby increasing the sample temperature. In order to verify the possibility that the RF pulses, and the resulting local temperature increase in HeLa cells, contribute the increase of cytosolic Ca^{2+} concentration, we therefore modified the protocol and left in-cell NMR samples inside the magnet for 2 hours without applying any RF pulses

before starting NMR measurements (Fig. 7.4.7-a).

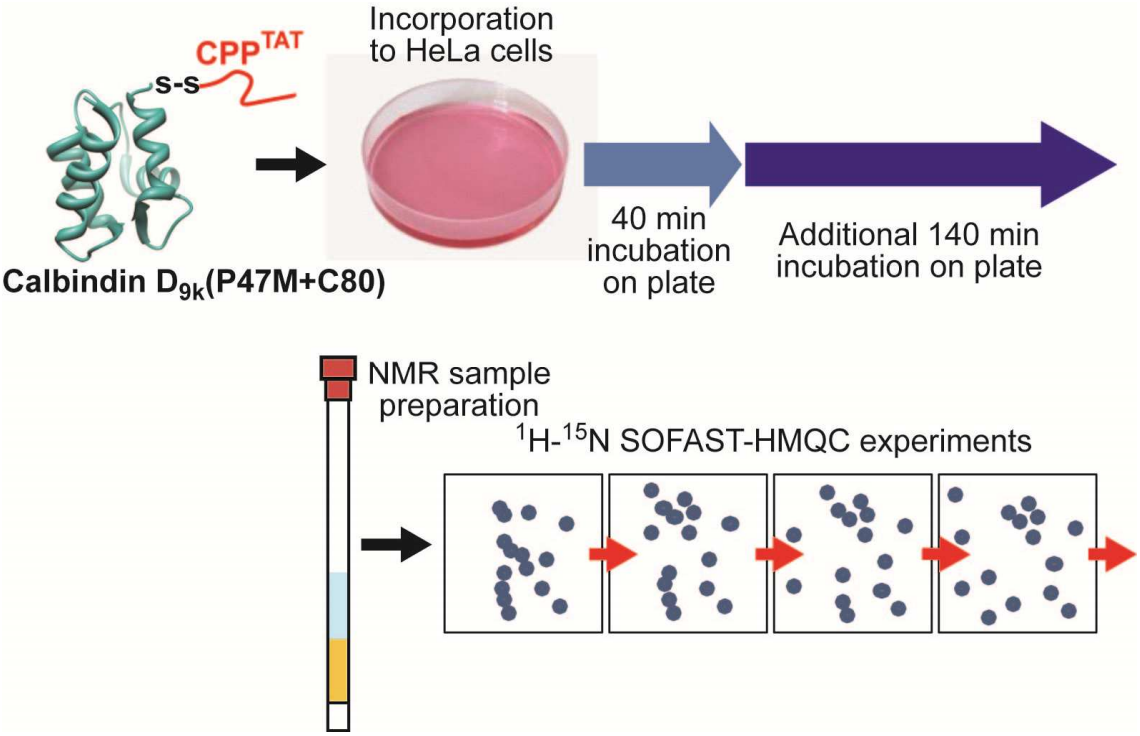


Fig.7.4.5

Schematic illustration of the in-cell NMR experiments of calbindin D_{9k}(P47M+C80) in HeLa cells with additional 140 minutes incubation on culture plates.

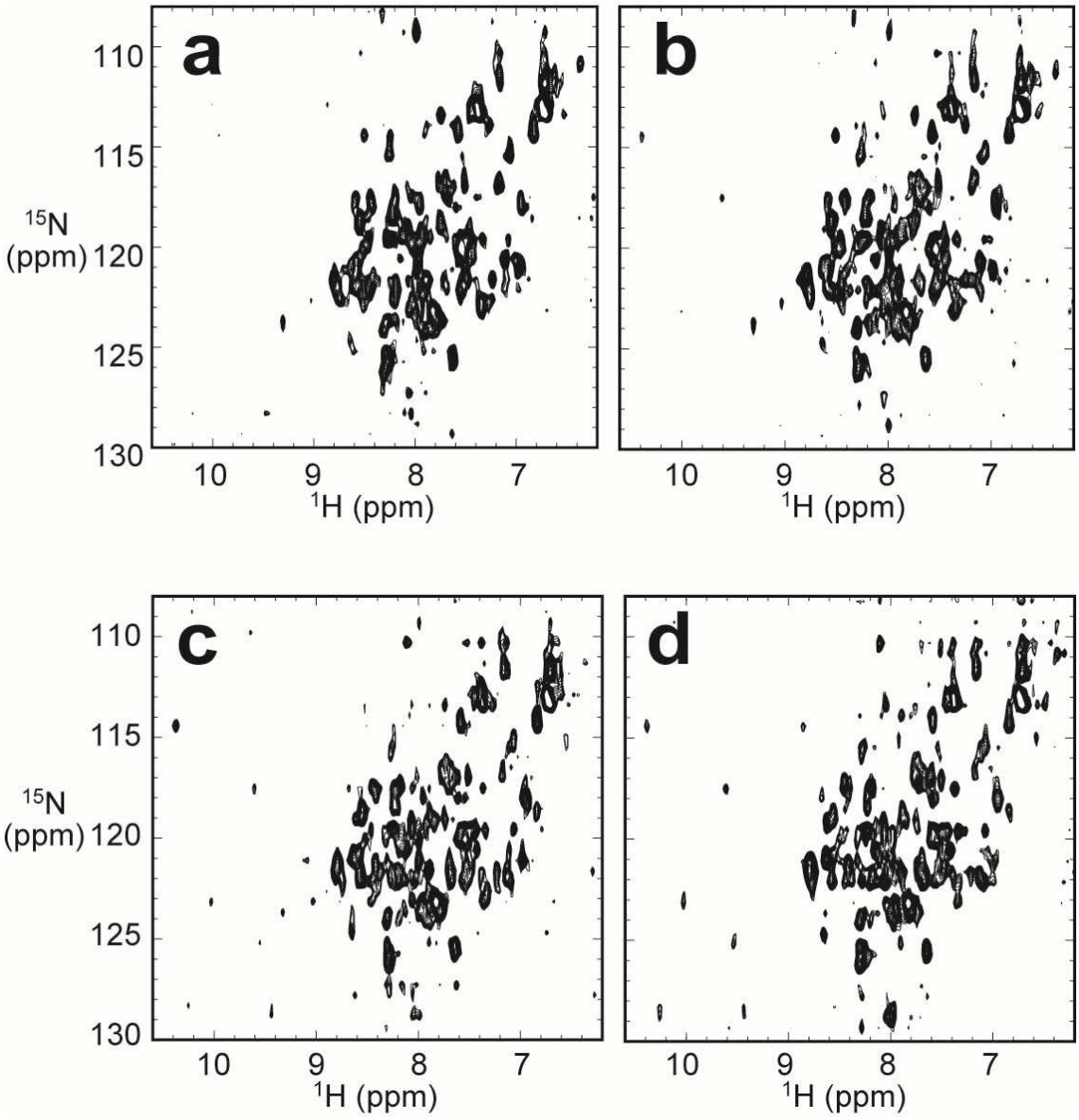


Fig.7.4.6

Time course of the 2D ^{1}H-^{15}N SOFAST-HMQC spectra of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells obtained by employing the modified protocol shown in Fig.7.4.5. The spectra were acquired during the period of 0 to 38 minutes (**a**), 39 to 76 minutes (**b**), 77 to 114 minutes (**c**), and 115 to 152 minutes (**d**) from the start of the in-cell NMR experiments.

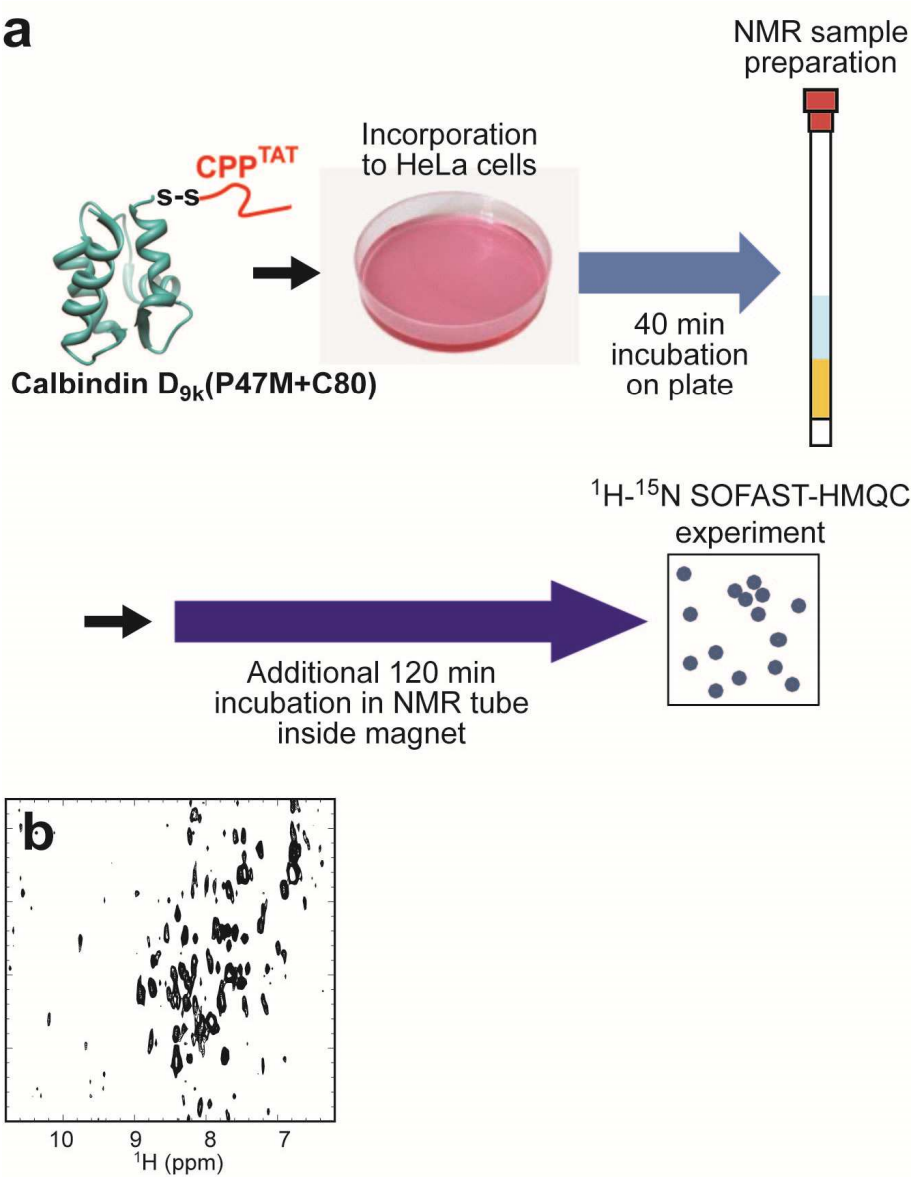


Fig.7.4.7

a. Schematic illustration of the in-cell NMR experiments of calbindin D_{9k}(P47M+C80) in HeLa cells with additional 120 minutes incubation in NMR tube inside magnet. **b**. The 2D ^{1}H-^{15}N SOFAST-HMQC spectrum of uniformly ^{15}N-labelled calbindin D_{9k}(P47M+C80) in HeLa cells obtained by employing the modified protocol shown in panel **a**.

The ^{1}H-^{15}N SOFAST-HMQC spectrum measured after this additional incubation time in the magnet showed essentially identical cross peak pattern (Fig.7.4.7-b) to that of the Ca^{2+}-bound state, suggesting that the NMR RF pulses is not the major cause for the increase of cytosolic Ca^{2+} concentration.

We have established a method and successfully monitor the increase of cytosolic Ca^{2+} concentration in HeLa cells by observing the time-dependent changes in the 2D ^{1}H-^{15}N correlation cross peaks of calbindin D_{9k}(P47M+C80), which was incorporated into the cells. The spectra from the metal free calboindin D_{9k} (P47M+C80) incorporated in to HeLa cells and the spectra from the Ca^{2+}-bound are similar, hence confirming that the metal bound to the protein released during the incorporation process. Though it is hard to predict properly, but the initial spectrum matches the Mg^{2+} bound *in vitro* spectrum, showing a clear indication that the protein is initially in the Mg^{2+}-bound state inside the HeLa cells. Then it was revealed that continuous increase of cytosolic Ca^{2+} concentration, which is detected as the time-dependent change in the in-cell NMR spectra of calbindin D_{9k}(P47M+C80) from the Mg^{2+}- to Ca^{2+}-bound states, is autonomously induced as responses to various stresses caused by removing cells from the culture dish by trypsin/EDTA and packed into an NMR tube, and the subsequently happened cell precipitation due to gravity, the starvation of oxygen and nutrient contents etc (Fig. 7.4.8). The RF pulses applied to HeLa cells during the NMR experiments are proved not to be the major source of this process.

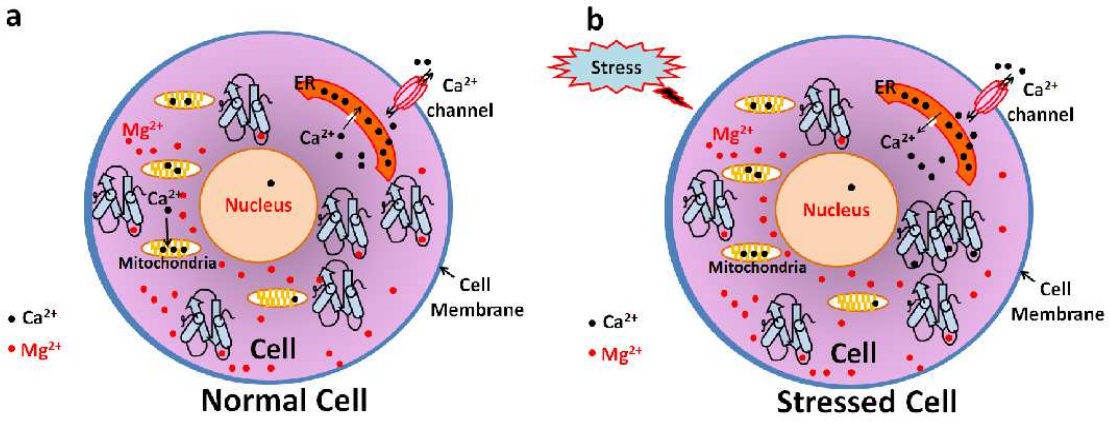


Fig.7.4.8

In normal cells (**a**) in order to maintain the Ca^{2+} level in cytosol, the Ca^{2+} are pumped into the ER or sometimes to mitochondria. The cells in stressed conditions (**b**), the ER releases Ca^{2+} to cytosol.

Chapter 8

Conclusion and Future prospects

8.1. Conclusion

I have optimised the protocol for the in-cell NMR method in HeLa cells and Ca^{2+} binding process inside the HeLa cells.

Optimisation for the incorporation of human calbindin D_{9k} into HeLa cells was successfully done. It was also optimised that the CPP^{TAT} bound target protein undergo conformational changes during membrane transportation as we have noticed that the Ca^{2+} was released during incorporation process. This incorporation process may not be a good idea for proteins with metal-binding activities.

Inside NMR tube, the HeLa cells were eventually stressed, and hence Ca^{2+} was released from ER, which was absorbed by calbindin D_{9k} protein. Though proteins' interaction with metals or other substrates inside cells have been studied by other biophysical methods, to the best of our knowledge, the results presented in this thesis are the first demonstration by in-cell NMR method, detecting cellular functions like releasing of Ca^{2+} due

to stress and further binding of these Ca^{2+} to the target protein inside cells in NMR time scale. These kinds of experiments can be tested with the cellular functions of various target protein under stressed condition inside cells. In this context, our method will provide a very useful tool for *in situ* monitoring of the "healthiness" of the cells in various in-cell NMR studies.

8.2. Future Prospects

Protein structure determination and its interaction with its partner molecule inside prokaryotic cells has been tested, and extending to mammalian cells is the next to be investigated. Short measurement time in terms of NMR pulse sequence like SOFAST, non-linear sampling techniques, etc. need to be optimized to acquire a quality NMR spectra in short time.

As was mentioned above one of the major limitations of the in-cell NMR experiments is the short life time of cells during the NMR measurements, and improvements of NMR apparatus, aiming at keeping cells alive by e.g. recycling the media during NMR measurement82, have been reported. In the future, in-cell NMR studies in mammalian cells may head in the direction of achieving more "physiological" condition during the experiments.

The other challenge is its inherent low sensitivity, making relatively high concentrations of the observed macromolecules necessary. In

particular, in-cell NMR investigations that focus on the observation of the behaviour of a particular protein in its natural environment are affected by this disadvantage. Most proteins occur in the cellular environment at the low μM to nM level. Interaction studies of an overexpressed protein that reached mM concentrations in the bacteria or mammalian cells with its natural interaction partners is, therefore, not exactly the physiological condition hence in most cases the current sensitivity of NMR spectrometers or methods in NMR spectrometer need to be enhanced.

In the method developing stage of mammalian in-cell NMR, the goals are to observe protein behaviours while cells are "alive" during the experiments and the physical and chemical properties of the targeted macromolecules. It is true that the cells have higher degree of selectivity for producing such biological macromolecule and they are not inert inside cells and having proper functions or dysfunction (leads to diseases). Keeping this type of scenario, in-cell NMR methods need to be carried out. Once these approaches are getting matured, more detailed information about the biological molecules inside cells and the condition of host cells, together called cellular structural biology can be investigated by in-cell NMR spectroscopy.

Publication

Hembram, D. S.S, Haremaki, T., Hamatsu, J., Inoue, J., Kamoshida, H., Ikeya, T., Mishima, M., Mikawa, T., Hayashi, N., Shirakawa, M. and Ito, Y. "An in-cell NMR study of monitoring stress-induced increase of cytosolic Ca^{2+} concentration in HeLa cells". *Biochem. Biophys. Res. Commun.* 2013 Aug 8. pii: S0006-291X(13)01312-0. doi: 10.1016/j.bbrc.2013.07.127. [Epub ahead of print]

Reference

1. Ellis, R.J., Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences* **26** (10), 597-604 (2001).

2. Minton, A.P., Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers* **20** (10), 2093-2120 (1981).

3. Miklos, A.C., Li, C., Sharaf, N.G., & Pielak, G.J., Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry* **49** (33), 6984-6991 (2010).

4. Charlton, L.M. *et al.*, Residue-level interrogation of macromolecular crowding effects on protein stability. *J Am Chem Soc* **130** (21), 6826-6830 (2008).

5. Hong, J. & Gierasch, L.M., Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J Am Chem Soc* **132** (30), 10445-10452 (2010).

6. Elcock, A.H., Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr Opin Struct Biol* **20** (2), 196-206 (2010).

7. Harada, R., Sugita, Y., & Feig, M., Protein crowding affects hydration structure and dynamics. *J Am Chem Soc* **134** (10), 4842-4849 (2012).

8. Lopes, M., Electron microscopy methods for studying in vivo DNA

replication intermediates. *Methods Mol Biol* **521**, 605-631 (2009).

9. Lucic, V., Rigort, A., & Baumeister, W., Cryo-electron tomography: The challenge of doing structural biology in situ. *J Cell Biol* **202** (3), 407-419.

10. Periasamy, A., Fluorescence resonance energy transfer microscopy: a mini review. *J Biomed Opt* **6** (3), 287-291 (2001).

11. van Roessel, P. & Brand, A.H., Imaging into the future: visualizing gene expression and protein interactions with fluorescent proteins. *Nat Cell Biol* **4** (1), E15-20 (2002).

12. Truong, K. *et al.*, FRET-based in vivo Ca2+ imaging by a new calmodulin-GFP fusion molecule. *Nat Struct Biol* **8** (12), 1069-1073 (2001).

13. Bachert, P., Pharmacokinetics using fluorine NMR in vivo. *Progress in Nuclear Magnetic Resonance Spectroscopy* **33** (1), 1-56 (1998).

14. Cohen, J.S., Lyon, R.C., & Daly, P.F., Monitoring intracellular metabolism by nuclear magnetic resonance. *Methods Enzymol* **177**, 435-452 (1989).

15. Gillies, R.J., *NMR in Physiology and Biomedicine*. (Academic Press, San Diego, 1994).

16. Serber, Z. & Dotsch, V., In-cell NMR spectroscopy. *Biochemistry* **40** (48), 14317-14323 (2001).

17. Serber, Z. *et al.*, High-resolution macromolecular NMR spectroscopy inside living cells. *Journal of the American Chemical Society* **123** (10), 2446-2447 (2001).

18. Bertrand, K., Reverdatto, S., Burz, D.S., Zitomer, R., & Shekhtman, A., Structure of proteins in eukaryotic compartments. *J Am Chem Soc* **134** (30), 12798-12806 (2012).

19. Banci, L. *et al.*, Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nat Chem Biol* **9** (5), 297-299 (2013).

20. Selenko, P., Serber, Z., Gade, B., Ruderman, J., & Wagner, G., Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. *Proceedings of the National Academy of Sciences of the United States of America* **103** (32), 11904-11909 (2006).

21. Sakai, T. *et al.*, In-cell NMR spectroscopy of proteins inside *Xenopus laevis* oocytes. *Journal of Biomolecular NMR* **36** (3), 179-188 (2006).

22. Inomata, K. *et al.*, High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458** (7234), 106-109 (2009).

23. Ogino, S. *et al.*, Observation of NMR signals from proteins introduced into living Mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin o. *Journal of the American Chemical Society* **131** (31), 10834-10835 (2009).

24. Reckel, S., Hänsel, R., Löhr, F., & Dötsch, V., In-cell NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **51** (2), 91–101 (2007).

25. Ito, Y. & Selenko, P., Cellular structural biology. *Curr Opin Struct Biol* **20** (5), 640-648 (2010).

26. Maldonado, A.Y., Burz, D.S., & Shekhtman, A., In-cell NMR spectroscopy. *Prog Nucl Magn Reson Spectrosc* **59** (3), 197-212 (2011).

27. Serber, Z., Ledwidge, R., Miller, S.M., & Dötsch, V., Evaluation of parameters critical to observing proteins inside living Escherichia coli by in-cell NMR spectroscopy. *Journal of the American Chemical Society* **123** (37), 8895-8901 (2001).

28. Dedmon, M.M., Patel, C.N., Young, G.B., & Pielak, G.J., FlgM gains structure in living cells. *Proceedings of the National Academy of Sciences of the United States of America* **99** (20), 12681-12684 (2002).

29. Wang, Q., Zhuravleva, A., & Gierasch, L.M., Exploring weak, transient protein--protein interactions in crowded in vivo environments by in-cell nuclear magnetic resonance spectroscopy. *Biochemistry* **50** (43), 9225-9236 (2011).

30. Burz, D.S., Dutta, K., Cowburn, D., & Shekhtman, A., Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nature Methods* **3** (2), 91-93 (2006).

31. Arnesano, F. *et al.*, Probing the interaction of cisplatin with the human

copper chaperone Atox1 by solution and in-cell NMR spectroscopy. *J Am Chem Soc* **133** (45), 18361-18369 (2011).

32. Sakakibara, D. *et al.*, Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature* **458** (7234), 102-105 (2009).

33. Ikeya, T. *et al.*, NMR protein structure determination in living E. coli cells using nonlinear sampling. *Nature Protocols* **5** (6), 1051-1060 (2010).

34. Miklos, A.C., Sarkar, M., Wang, Y., & Pielak, G.J., Protein crowding tunes protein stability. *J Am Chem Soc* **133** (18), 7116-7120 (2011).

35. Waudby, C.A. *et al.*, Rapid distinction of intracellular and extracellular proteins using NMR diffusion measurements. *J Am Chem Soc* **134** (28), 11312-11315 (2012).

36. Selenko, P. *et al.*, In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol* **15** (3), 321-329 (2008).

37. Renaud, J.P. & Delsuc, M.A., Biophysical techniques for ligand screening and drug design. *Curr Opin Pharmacol* **9** (5), 622-628 (2009).

38. Campbell, A.K., *Intracellular Calcium: Its Universal Role as Regulator*. (John Wiley & Sons Ltd, New York, 1983).

39. Berridge, M.J. & Irvine, R.F., Inositol trisphosphate, a novel second

messenger in cellular signal transduction. *Nature* **312** (5992), 315-321 (1984).

40. Nishizuka, Y., Studies and perspectives of protein kinase C. *Science* **233** (4761), 305-312 (1986).

41. Schafer, B.W. & Heizmann, C.W., The S100 family of EF-hand calcium-binding proteins: functions and pathology. *Trends Biochem Sci* **21** (4), 134-140 (1996).

42. Kruman, I., Guo, Q., & Mattson, M.P., Calcium and reactive oxygen species mediate staurosporine-induced mitochondrial dysfunction and apoptosis in PC12 cells. *J Neurosci Res* **51** (3), 293-308 (1998).

43. Tombal, B., Denmeade, S.R., & Isaacs, J.T., Assessment and validation of a microinjection method for kinetic analysis of $[Ca^{2+}]i$ in individual cells undergoing apoptosis. *Cell Calcium* **25** (1), 19-28 (1999).

44. Pinton, P., Giorgi, C., Siviero, R., Zecchini, E., & Rizzuto, R., Calcium and apoptosis: ER-mitochondria Ca^{2+} transfer in the control of apoptosis. *Oncogene* **27** (50), 6407-6418 (2008).

45. Maler, L., Blankenship, J., Rance, M., & Chazin, W.J., Site-site communication in the EF-hand Ca^{2+}-binding protein calbindin D_{9k}. *Nat Struct Biol* **7** (3), 245-250 (2000).

46. Li, G. *et al.*, Role of ERO1-alpha-mediated stimulation of inositol 1,4,5-triphosphate receptor activity in endoplasmic reticulum stress-induced apoptosis. *J Cell Biol* **186** (6), 783-792 (2009).

47. Sim, J.Y., Jung, E.M., Yoo, Y.M., Choi, K.C., & Jeung, E.B., Transcriptional and translational expression of calbindin-D_{9k} in the duodenum, kidney and uterus of a female canine model. *J Vet Sci* **11** (1), 15-19 (2010).

48. Marenholz, I., Heizmann, C.W., & Fritz, G., S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochem Biophys Res Commun* **322** (4), 1111-1122 (2004).

49. Schwaller, B., Cytosolic Ca^{2+} buffers. *Cold Spring Harb Perspect Biol* **2** (11), a004051 (2011).

50. Chazin, W.J. *et al.*, Proline isomerism leads to multiple folded conformations of calbindin D_{9k}: direct evidence from two-dimensional ^{1}H NMR spectroscopy. *Proc Natl Acad Sci U S A* **86** (7), 2195-2198 (1989).

51. Wüthrich, K., *NMR of Proteins and Nucleic Acids*. (Wiley, New York, 1986).

52. Bax, A., Multidimensional nuclear magnetic resonance methods for protein studies. *Current Opinion in Structural Biology* **4** (5), 738-744 (1994).

53. Kay, L.E., Ikura, M., Tschudin, R., & Bax, A., Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* **89** (3), 496-514 (1990).

54. Barna, J.C.J., Laue, E.D., Mayger, M.R., Skilling, J., & Worrall, S.J.P., Exponential sampling, an alternative method for sampling in two-dimensional NMR experiments. *Journal of Magnetic Resonance* **73** (1), 69-77 (1987).

55. Schmieder, P., Stern, A.S., Wagner, G., & Hoch, J.C., Improved resolution in triple-resonance spectra by nonlinear sampling in the constant-time domain. *Journal of Biomolecular NMR* **4** (4), 483-490 (1994).

56. Schanda, P., Kupce, E., & Brutscher, B., SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *Journal of Biomolecular NMR* **33** (4), 199-211 (2005).

57. Schanda, P., Van Melckebeke, H., & Brutscher, B., Speeding up three-dimensional protein NMR experiments to a few minutes. *Journal of the American Chemical Society* **128** (28), 9042-9043 (2006).

58. Frydman, L., Scherf, T., & Lupulescu, A., The acquisition of multidimensional NMR spectra within a single scan. *Proc Natl Acad Sci U S A* **99** (25), 15858-15862 (2002).

59. Laue, E.D., Mayger, M.R., Skilling, J., & Staunton, J., Reconstruction of phase sensitive 2D NMR spectra by maximum entropy. *Journal of Magnetic Resonance* **68**, 14-29 (1986).

60. Hoch, J.C., Stern, A.S., Donoho, D., & Johnstone, I., Maximum entropy

reconstruction of complex (phase-sensitive) spectra. *J Magn Reson* **86**, 236-246 (1990).

61. Orekhov, V.Y., Ibraghimov, I.V., & Billeter, M., MUNIN: a new approach to multi-dimensional NMR spectra interpretation. *J Biomol NMR* **20** (1), 49-60 (2001).

62. Jaravine, V., Ibraghimov, I., & Orekhov, V.Y., Removal of a time barrier for high-resolution multidimensional NMR spectroscopy. *Nature Methods* **3** (8), 605-607 (2006).

63. Hyberts, S.G. *et al.*, Ultrahigh-resolution (1)H-(13)C HSQC spectra of metabolite mixtures using nonlinear sampling and forward maximum entropy reconstruction. *J Am Chem Soc* **129** (16), 5108-5116 (2007).

64. Kazimierczuk, K. & Orekhov, V.Y., Accelerated NMR spectroscopy by using compressed sensing. *Angew Chem Int Ed Engl* **50** (24), 5556-5559.

65. Holland, D.J., Bostock, M.J., Gladden, L.F., & Nietlispach, D., Fast multidimensional NMR spectroscopy using compressed sensing. *Angew Chem Int Ed Engl* **50** (29), 6548-6551 (2011).

66. Hamatsu, J. *et al.*, High-resolution heteronuclear multidimensional NMR of proteins in living insect cells using a baculovirus protein expression system. *J Am Chem Soc* **135** (5), 1688-1691 (2013).

67. Marion, D., Ikura, M., Tschudin, R., & Bax, A., Rapid recording of 2D NMR spectra without phase cycling. Application to the study of

hydrogen exchange in proteins. *J Magn Reson* **85** (2), 393-399 (1989).

68. Grzesiek, S. & Bax, A., Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* **114** (16), 6291-6293 (1992).

69. Grzesiek, S. & Bax, A., An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J Magn Reson* **99** (1), 201-207 (1992).

70. Piotto, M., Saudek, V., & Sklenar, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* **2** (6), 661-665 (1992).

71. Grzesiek, S. & Bax, A., Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* **96** (2), 432-440 (1992).

72. Clubb, R.T., Thanabal, V., & Wagner, G., A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{15}N, and $^{13}C'$ chemical shifts in ^{15}N-^{13}C-labelled proteins. *J Magn Reson* **97** (1), 213-217 (1992).

73. Shaka, A.J., Lee, C.J., & Pines, A., Iterative schemes for bilinear operators; application to spin decoupling. *J Magn Reson* **77** (2), 274-293 (1988).

74. Emsley, L. & Bodenhausen, G., Gaussian pulse cascades: New analytical functions for rectangular selective inversion and in-phase

excitation in NMR. *Chemical Physics Letters* **165** (6), 469-476 (1990).

75. Kupce, E., Boyd, J., & Campbell, I.D., Short Selective Pulses for Biochemical Applications. *J Magn Reson, Ser. B* **106** (3), 300-303 (1995).

76. McCoy, M.A. & Mueller, L., Selective shaped pulse decoupling in NMR: homonuclear [^{13}C]carbonyl decoupling. *J Am Chem Soc* **114** (6), 2108-2112 (1992).

77. Gull, S. & Skilling, J., *Quantied Maximum Entropy: MemSys5 Users' Manual*. (1999).

78. Vranken, W.F. *et al.*, The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59** (4), 687-696 (2005).

79. Kaczanowski, S. & Zielenkiewicz, P., Why similar protein sequences encode similar three-dimensional structures? *Theoretical Chemistry Accounts* **125**, 543-550 (2010).

80. Chothia, C. & Lesk, A.M., The relation between the divergence of sequence and structure in proteins. *EMBO J* **5** (4), 823-826 (1986).

81. Geen, H. & Freeman, R., Band-selective radiofrequency pulses. *J Magn Reson* **93** (1), 93-141 (1991).

82. Kubo, S. *et al.*, A gel-encapsulated bioreactor system for NMR studies of protein-protein interactions in living mammalian cells. *Angew Chem Int Ed Engl* **52** (4), 1208-1211 (2013).